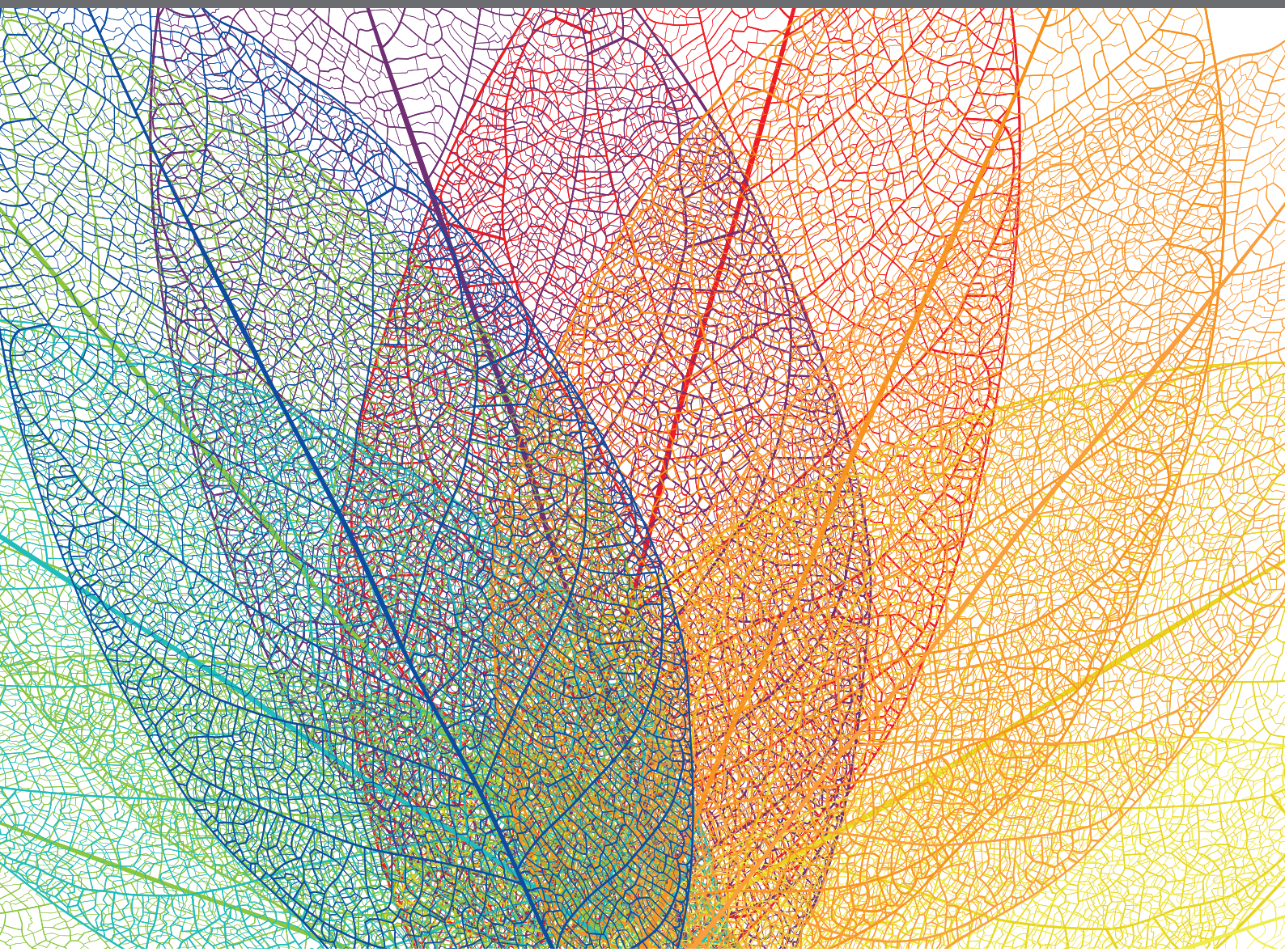
THE BRASSICACEAE — AGRI-HORTICULTURAL AND ENVIRONMENTAL PERSPECTIVES

EDITED BY: Naser A. Anjum, Sarvajeet Singh Gill, Om Parkash Dhankher, Juan F. Jimenez and Narendra Tuteja

PUBLISHED IN: Frontiers in Plant Science



frontiers Research Topics

Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88945-645-1
DOI 10.3389/978-2-88945-645-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

THE BRASSICACEAE — AGRI-HORTICULTURAL AND ENVIRONMENTAL PERSPECTIVES

Topic Editors:
Naser A. Anjum, Aligarh Muslim University, India
Sarvajeet Singh Gill, Maharshi Dayanand University, India
Om Parkash Dhankher, University of Massachusetts Amherst, United States
Juan F. Jimenez, Institute for Scientific and Technological Research, Mexico
Narendra Tuteja, International Center for Genetic Engineering and Biotechnology, India

This Frontiers Research Topic "The Brassicaceae- Agri-Horticultural and Environmental Perspectives" is an effort to provide a common platform to agronomists, horticulturists, plant breeders, plant geneticists/molecular biologists, plant physiologists and environmental plant scientists exploring major insights into the role of important members of the plant family Brassicaceae (the mustard family, or Cruciferae) in agri-horticultural and environmental arenas.

Citation: Anjum, N. A., Gill, S. S., Dhankher, O. P., Jimenez, J. F., Tuteja, N., eds. (2018). The Brassicaceae — Agri-Horticultural and Environmental Perspectives. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-645-1

Table of Contents

06 ***Editorial: The Brassicaceae—Agri-Horticultural and Environmental Perspectives***
Naser A. Anjum, Sarvajeet S. Gill, Om P. Dhankher, Juan F. Jimenez and Narendra Tuteja

I. YIELD AND NUTRITIONAL QUALITY

10 ***Floral Initiation in Response to Planting Date Reveals the Key Role of Floral Meristem Differentiation Prior to Budding in Canola (*Brassica napus *L.)***
Yaofeng Zhang, Dongqing Zhang, Huasheng Yu, Baogang Lin, Ying Fu and Shuijin Hua

24 ***High-Density SNP Map Construction and QTL Identification for the Apetalous Character in* Brassica napus *L.***
Xiaodong Wang, Kunjiang Yu, Hongge Li, Qi Peng, Feng Chen, Wei Zhang, Song Chen, Maolong Hu and Jiefu Zhang

39 ***Genome-Wide Identification of QTL for Seed Yield and Yield-Related Traits and Construction of a High-Density Consensus Map for QTL Comparison in* Brassica napus**
Weiguo Zhao, Xiaodong Wang, Hao Wang, Jianhua Tian, Baojun Li, Li Chen, Hongbo Chao, Yan Long, Jun Xiang, Jianping Gan, Wusheng Liang and Maoteng Li

53 ***Multigenic Control of Pod Shattering Resistance in Chinese Rapeseed Germplasm Revealed by Genome-Wide Association and Linkage Analyses***
Jia Liu, Jun Wang, Hui Wang, Wenxiang Wang, Rijin Zhou, Desheng Mei, Hongtao Cheng, Juan Yang, Harsh Raman and Qiong Hu

67 ***Identification of microRNAs and Their Target Genes Explores miRNA-Mediated Regulatory Network of Cytoplasmic Male Sterility Occurrence during Anther Development in Radish (*Raphanus sativus L.*)***
Wei Zhang, Yang Xie, Liang Xu, Yan Wang, Xianwen Zhu, Ronghua Wang, Yang Zhang, Everlyne M. Muleke and Liwang Liu

83 ***Transcriptomic Analysis Identifies Differentially Expressed Genes (DEGs) Associated With Bolting and Flowering in Radish (*Raphanus sativus L.*)***
Shanshan Nie, Chao Li, Yan Wang, Liang Xu, Everlyne M. Muleke, Mingjia Tang, Xiaochuan Sun and Liwang Liu

95 **De novo *Taproot Transcriptome Sequencing and Analysis of Major Genes Involved in Sucrose Metabolism in Radish (*Raphanus Sativus L.*)***
Rugang Yu, Liang Xu, Wei Zhang, Yan Wang, Xiaobo Luo, Ronghua Wang, Xianwen Zhu, Yang Xie, Benard Karanja and Liwang Liu

107 ***Modification of Leaf Glucosinolate Contents in* Brassica Oleracea *by Divergent Selection and Effect on Expression of Genes Controlling Glucosinolate Pathway***
Tamara Sotelo, Pablo Velasco, Pilar Soengas, Víctor M. Rodríguez and María E. Cartea

119 ***Plants as Biofactories: Postharvest Stress-Induced Accumulation of Phenolic Compounds and Glucosinolates in Broccoli Subjected to Wounding Stress and Exogenous Phytohormones***
Daniel Villarreal-García, Vimal Nair, Luis Cisneros-Zevallos and Daniel A. Jacobo-Velázquez

130 ***Characterizing Variation of Branch Angle and Genome-Wide Association Mapping in Rapeseed (*Brassica napus *L.)***
Jia Liu, Wenxiang Wang, Desheng Mei, Hui Wang, Li Fu, Daoming Liu, Yunchang Li and Qiong Hu

140 ***Intraspecific Variability of Floral Nectar Volume and Composition in Rapeseed (*Brassica napus *L. var.* Oleifera*)***
Michele Bertazzini and Giuseppe Forlani

II. BREEDING STUDY-OUTCOMES

153 ***Genome-Wide Single-Nucleotide Polymorphisms Discovery and High-Density Genetic Map Construction in Cauliflower Using Specific-Locus Amplified Fragment Sequencing***
Zhenqing Zhao, Honghui Gu, Xiaoguang Sheng, Huifang Yu, Jiansheng Wang, Long Huang and Dan Wang

162 ***Genetic and Epigenetic Alterations of* Brassica nigra *Introgression Lines From Somatic Hybridization: A Resource for Cauliflower Improvement***
Gui-xiang Wang, Jing Lv, Jie Zhang, Shuo Han, Mei Zong, Ning Guo, Xing-ying Zeng, Yue-yun Zhang, You-ping Wang and Fan Liu

174 ***Genome-Specific Differential Gene Expressions in Resynthesized* Brassica *Allotetraploids From Pair-Wise Crosses of Three Cultivated Diploids Revealed by RNA-seq***
Dawei Zhang, Qi Pan, Cheng Cui, Chen Tan, Xianhong Ge, Yujiao Shao and Zaiyun Li

187 ***The miRNAs and Their Regulatory Networks Responsible for Pollen Abortion in Ogura-CMS Chinese Cabbage Revealed by High-Throughput Sequencing of miRNAs, Degradomes, and Transcriptomes***
Xiaochun Wei, Xiaohui Zhang, Qiuju Yao, Yuxiang Yuan, Xixiang Li, Fang Wei, Yanyan Zhao, Qiang Zhang, Zhiyong Wang, Wusheng Jiang and Xiaowei Zhang

202 ***Cytogenetic Diversity of Simple Sequences Repeats in Morphotypes of* Brassica Rapa *ssp.* Chinensis**
Jin-shuang Zheng, Cheng-zhen Sun, Shu-ning Zhang, Xi-lin Hou and Guusje Bonnema

212 ***Genome-Wide DNA Methylation Profiling by Modified Reduced Representation Bisulfite Sequencing in* Brassica Rapa *Suggests That Epigenetic Modifications Play a Key Role in Polyploid Genome Evolution***
Xun Chen, Xianhong Ge, Jing Wang, Chen Tan, Graham J. King and Kede Liu

224 ***Comparative Leaves Transcriptome Analysis Emphasizing on Accumulation of Anthocyanins in* Brassica*: Molecular Regulation and Potential Interaction With Photosynthesis***
Muhammad A. Mushtaq, Qi Pan, Daozong Chen, Qinghua Zhang, Xianhong Ge and Zaiyun Li

239 ***Chromosome Doubling of Microspore-Derived Plants From Cabbage (*Brassica Oleracea *var.* Capitata *L.) and Broccoli (*Brassica Oleracea *var.* Italica *L.)***
Suxia Yuan, Yanbin Su, Yumei Liu, Zhansheng Li, Zhiyuan Fang, Limei Yang, Mu Zhuang, Yangyong Zhang, Honghao Lv and Peitian Sun

249 ***An Efficient Method for Adventitious Root Induction From Stem Segments of* Brassica *Species***
Sandhya Srikanth, Tsui Wei Choong, An Yan, Jie He and Zhong Chen

III. STRESS MANAGEMENT AND MINERAL NUTRITION

260 ***Molecular Breeding in* Brassica *for salt Tolerance: Importance of Microsatellite (SSR) Markers for Molecular Breeding in* Brassica**
Manu Kumar, Ju-Young Choi, Nisha Kumari, Ashwani Pareek and Seong-Ryong Kim

273 **Arabidopsis AtDjA3 *Null Mutant Shows Increased Sensitivity to Abscisic Acid, Salt, and Osmotic Stress in Germination and Post-germination Stages***
Silvia Salas-Muñoz, Aída A. Rodríguez-Hernández, Maria A. Ortega-Amaro, Fatima B. Salazar-Badillo and Juan F. Jiménez-Bremont

284 ***Genome-Wide Analysis of the Glutathione* S*-Transferase Gene Family in* Capsella Rubella*: Identification, Expression, and Biochemical Functions***
Gang He, Chao-Nan Guan, Qiang-Xin Chen, Xiao-Jun Gou, Wei Liu, Qing-Yin Zeng and Ting Lan

298 ***Nitrogen-Efficient and Nitrogen-Inefficient Indian Mustard Showed Differential Expression Pattern of Proteins in Response to Elevated CO_2 and Low Nitrogen***
Peerzada Y. Yousuf, Arshid H. Ganie, Ishrat Khan, Mohammad I. Qureshi, Mohamed M. Ibrahim, Maryam Sarwat, Muhammad Iqbal and Altaf Ahmad

320 ***Interactions of Sulfate With Other Nutrients As Revealed by H2S Fumigation of Chinese Cabbage***
Martin Reich, Muhammad Shahbaz, Dharmendra H. Prajapati, Saroj Parmar, Malcolm J. Hawkesford and Luit J. De Kok

328 **Brassica Napus *Genome Possesses Extraordinary High Number of* CAMTA *Genes and* CAMTA3 *Contributes to PAMP Triggered Immunity and Resistance to* Sclerotinia sclerotiorum**
Hafizur Rahman, You-Ping Xu, Xuan-Rui Zhang and Xin-Zhong Cai

IV. ENVIRONMENTAL PERSPECTIVES

344 ***Growth and Metal Accumulation of an* Alyssum Murale *Nickel Hyperaccumulator Ecotype Co-cropped With* Alyssum Montanum *and Perennial Ryegrass in Serpentine Soil***
Catherine L. Broadhurst and Rufus L. Chaney

353 **De novo *Transcriptome Analysis of* Sinapis Alba *in Revealing the Glucosinolate and Phytochelatin Pathways***
Xiaohui Zhang, Tongjin Liu, Mengmeng Duan, Jiangping Song and Xixiang Li

frontiers in Plant Science

EDITORIAL
published: 27 August 2018
doi: 10.3389/fpls.2018.01141

Editorial: The Brassicaceae—Agri-Horticultural and Environmental Perspectives

Naser A. Anjum[1†], Sarvajeet S. Gill[2*†], Om P. Dhankher[3†], Juan F. Jimenez[4†] and Narendra Tuteja[5*†]*

[1] CESAM-Centre for Environmental and Marine Studies, University of Aveiro, Aveiro, Portugal, [2] Stress Physiology and Molecular Biology Lab, Centre for Biotechnology, MD University, Rohtak, India, [3] Stockbridge School of Agriculture, University of Massachusetts Amherst, Amherst, MA, United States, [4] Institute for Scientific and Technological Research, San Luis Potosí, Mexico, [5] Plant Molecular Biology Group, International Center for Genetic Engineering and Biotechnology, New Delhi, India

Keywords: brassicaceae family, agricultural-horticultural perspective, environmental health, plant breeding, crop improvement

Editorial on the Research Topic

The Brassicaceae—Agri-Horticultural and Environmental Perspectives

OPEN ACCESS

Edited and reviewed by:
Marcelino Perez De La Vega, Universidad de León, Spain

***Correspondence:**
Naser A. Anjum
anjum@ua.pt
Sarvajeet S. Gill
ssgill14@yahoo.co.in
Narendra Tuteja
ntuteja@icgeb.res.in

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *27 February 2018*
Accepted: *13 July 2018*
Published: *27 August 2018*

Citation:
Anjum NA, Gill SS, Dhankher OP, Jimenez JF and Tuteja N (2018) Editorial: The Brassicaceae—Agri-Horticultural and Environmental Perspectives. Front. Plant Sci. 9:1141. doi: 10.3389/fpls.2018.01141

Human health is closely related with the environmental health; therefore, efforts are being made to improve their interrelationship. In this context, a number of plant genera have been identified which can significantly contribute to the protection of the environment as well as human health. Notably, Brassicaceae is one of the prominent plant families with *Arabidopsis*, *Alyssum,* and *Brassica* as model plants; *Boechera*, *Brassica*, and *Cardamine* as developing model generic systems; and radish, rocket, watercress, wasabi, horseradish, vegetable, and oil crops as cultivated plant species. Thus, most members of Brassicaceae have been the subjects of exhaustive research mainly due to their innumerable multidimensional roles in the current context. The Frontiers Research Topic "*The Brassicaceae—Agri-Horticultural and Environmental Perspectives*" aims to highlight major current global research outcomes in this direction. This research topic incorporated 28 publications including 27 research papers and one review article. Molecular-genetic insights into agri-horticultural aspects were dealt with in 20 of the publications, six reports were on the stress management and nutrition aspects and two reports on environmental perspectives.

YIELD AND NUTRITIONAL QUALITY

Improving yield and nutritional quality has been an important focus of the agri-horticultural studies on the Brassicaceae family members. The formations of the flower and siliques directly affect seed yield (SY); hence they are considered important for *B. napus* production, where delayed planting can impact SY. Additionally, the period from floral meristem differentiation to budding governs effective flower and siliques formation (Zhang et al.). In this context, maximizing effective flower numbers prior to budding and reducing the degradation of the floral meristem can improve silique numbers and its final formation, respectively. *Brassica* oilseed genotypes with apetalous flowers are a component of the high-yielding ideotype and also of particular interest in breeding programs. *B. napus* possess nine consensus quantitative trait loci (QTLs) present on the A3, A5, A6, A9, and C8 chromosomes (Wang et al.). Notably, the population with QTLs *qPD.C8-2* and *qPD.C8-3* can exhibit a high heritability of petalous degree (PDgr) and less sensitivity to environment. In another report, among seven major QTLs identified by Zhao et al., three and two QTLs were for SY and seed weight (SW), respectively; whereas, one QTL each was for branch height (BH) and branch number (FBN).

Particularly, the expression of QTLs for SY can occur stably in winter (*cqSY-C6-2* and *cqSY-C6-3*) or in spring within the rapeseed cultivation area (*cqSY-A2-2*). Seed yield losses due to shatter have been a great issue in commercial harvesting of *Brassica* genotypes. In particular, *B. napus* exhibits a multigenic inheritance for pod shatter resistance. In a diversity panel, doubled haploid and intermated possess six QTLs for resistance to pod shatter (Liu et al.). These QTLs are located on chromosomes A01, A06, A07, A09, C02, and C05, where QTL *qSRI.A09* and *qSRI.A06* can occur across environments.

In radish (*Rhaphanus sativus*), information is scanty on the occurrence of cytoplasmic male sterility (CMS; a maternally inherited trait incapable of producing functional pollen) at the posttranscriptional level. *R. sativus* male sterile line "WA" and its maintainer line "WB" were revealed to involve a potential miRNA-mediated regulatory network of CMS during anther development (Zhang et al.). In *R. sativus*, premature bolting results in poor root growth and reduced harvest. Differentially expressed genes (DEGs) related to *R. sativus* bolting and flowering involve the model of 24 miRNA–DEG pairs (Nie et al.). Particularly, the intricate genetic networks of bolting and flowering in *R. sativus* involve the pairs including miR5227–VRN1, miR6273–PRP39, and miR860–NF-YB3. Notably, the major molecular mechanism underlying the complex *R. sativus* taproot development process is in its infancy. *De novo* taproot transcriptome sequencing and analysis of major genes involved in sucrose metabolism discovered a total of 103 unigenes, encoding eight enzymes involved in the sucrose metabolism-related pathways (Yu et al.).

It is possible to modify the content of secondary metabolites, including glucosinolates (GSLs), and obtain vegetables enriched in these compounds, which are related to plant defense and human health. The developed six *Brassica oleracea* var. *acephala* genotypes exhibit high and low contents of three major GSLs namely sinigrin (SIN), glucoiberin (GIB), and glucobrassicin (GBS) (Sotelo et al.). Compared with the divergent selection, the use of mass selection can be an efficient way of modifying the SIN and GIB concentrations that are related with the GSL–ALK locus that, in turn, varies with the *CYP81F2* gene expression. Storing the whole heads of broccoli (*B. oleracea* var. italica) at 20°C for 24 h exhibits several specific phenolic compounds such as 1,2,2-trisinapoylgentiobiose (2,2-TSG), 3-O-Caffeoylquinic acids (3-O-CQA), 1,2-disinapoylgentiobiose (1,2-DSG), 1,2-diferulolylgentiobiose (1,2-DFG) and 1,2-disinapoyl-2-ferulolylgentiobiose (1,2-DS-2-FG) (Villarreal-García et al.). Additionally, the tissue level of these phenolic compounds as well as GSLs can also be modulated by wounding stress alone and in combination with exogenous phytohormones including methyl jasmonate or ethylene.

As a plant architectural trait, branch angle is the basic requirement for high-density cultivation and mechanical harvesting in rapeseed (*Brassica napus*). In a panel of 143 elite *B. napus* accessions, the 60 K Illumina Infinium Single Nucleotide Polymorphisms array analyses revealed significant natural phenotypic variations in branch angle (Liu et al.). Four QTLs for branch angle included two Lazy (AT5G14090) orthologous, SPL14 (AT1G20980) and auxin-responsive GH3 family protein (AT5G51470), genes. Rapeseed possesses significant melliferous potentials and represents a main forage crop for bees. In winter genotypes of field-grown *B. napus* var. oleifera, the volume of the nectar, a nutrient-rich aqueous solution, can vary and occur in the range of 0.02–0.75 μL flower^{-1} (Bertazzini and Forlani). Additionally, the phloem sap composition is modulated by the nectar-variability that, in turn, depends on nectary metabolism.

BREEDING STUDY OUTCOMES

Specific locus amplified fragment sequencing (SLAFS) can be utilized to compare genomic studies within the *B. oleracea* species, QTL identification, and their molecular breeding. In a double-haploid, segregating population of *B. oleracea*, the genetic linkage map constructed with SLAFS revealed a total genetic length of 890.01 cM (with an average marker interval of 0.50 cM), and covered 364.9 Mb of the reference genome (Zhao et al.). Asymmetric somatic hybridization and alien DNA introgression can contribute to the genetic and epigenetic alterations of somatically hybridized black mustard "G1/1" (*B. nigra*, $2n = 16$, BB genome) introgression lines and can also be a major resource for breeding *B. oleracea* var. *botrytis* ($2n = 18$, CC genome) (Wang et al.). Global transcriptome analyses through RNA-Seq confirmed the expression of transgressive gene in *Brassica* allotetraploids (*B. juncea*, AABB; *B. napus*, AACC; *B. carinata*, BBCC) as genome-wide and temporal, where transgressive up-regulation of resistance-related genes actually controls their immediate physiological pre-adaptation (Zhang et al.). Therein, the silique walls can exhibit much wider variations than leaves after the onset of the tissue-specific expression partitioning quickly after the allotetraploids formation and is associated with the expression of r-protein genes. Notably, the maturity of silique walls and exhibition therein of low translation activity (vs. young leaves) can control the exhibition of expression levels of the r-protein gene in silique lower than in leaves. This can be further corroborated with the expression of the rRNA genes (silenced in vegetative tissues) in reproductive tissues such as sepals and petals. Information is meager on the microRNAs (miRNAs) underlying flower bud development and their potential response to the Ogura-cytoplasmic male sterility (Ogura-CMS), a CMS type with complete male sterility and stability in Chinese cabbage (*B. rapa* ssp. *pekinensis*). *B. rapa* ssp. *pekinensis* buds from both Ogura-CMS and its maintainer possess 426 novel miRNAs, where a regulatory network involving two novel miRNA/target cascades (novel-miR-335/H+-ATPase and novel-miR-448/SUC1) contributes to bud development, especially for pollen engenderation (Wei et al.). Cytogenetic diversity of simple sequence repeats (SSRs) among morphotypes of *B. rapa* ssp. *chinensis* can be revealed by fluorescence *in situ* hybridization. Besides, this method can also confirm nonrandom and motif-dependent chromosomal locations of mono-, di-, and tri-nucleotide repeat loci (Zheng et al.). In fact, differences between SSR repeats with respect to abundance and distribution lead to the genomic evolution of *B. rapa* species. In order to profile the genome-wide DNA methylation, the available methods

so far, such as whole-genome bisulfite sequencing (including MethylC-seq and BS-seq), perform single-cytosine methylation resolution and directly estimate the proportion of molecules methylated. Whereas, enrichment of CG-rich sequences in the genome can be possible with the reduced representation bisulfite sequencing (RRBS) method. A new double enzyme-digested RRBS method exhibits a consistent percentage of CG, CHG, and CHH loci located in genic regions between enriched targeted regions and can help in genome-wide DNA methylation profiling at single-base resolution in plants such as *B. rapa* (Chen et al.). The modified RRBS can be a cost-effective, simple, and suitable method for DNA methylation-profiling of large natural populations or for construction of DNA methylation genetic maps.

There occurs a crosstalk between anthocyanin and photosynthesis in *Brassica* species with purple or green leaves. Therein, the upregulation of both transcription factors (TFs), including *Transparent Testa* 8 (*TT8*), and *Transparent Testa 19* (*TT19*), and the anthocyanin late biosynthetic genes (LBGs), especially *dihydroflavonol 4-reductase* (*DFR*) and *anthocyanidin synthase* (*ANS*), contributes to anthocyanin production (Mushtaq et al.). Notably, the upregulation of cytosolic 6-phosphogluconolactonase (*PLG5*) that is involved in the oxidative pentose phosphate pathway points toward the involvement of photosynthesis in the purple color of leaves. Additionally, downregulation of three genes *FTSH PROTEASE 8* (*FTS8*), *GLYCOLATE OXIDASE 1* (*GOX1*), and *GLUTAMINE SYNTHETASE 1;4* (*GLN1;4*) related with the degradation of photo-damaged proteins in photosystem II and light respiration also corroborates the photosynthesis role in the purple color of leaves. In 29 populations of microspore-derived plantlets from cabbage (*B. oleracea* var. *capitata*) and broccoli (*B. oleracea* var. *italica*), the chromosome doubling can be random and genotype-dependent (Yuan et al.). Compared with the result of the total spontaneous doubling of 0 to 76.9% and 52.2 to 100% in 14 *B. oleracea* var. *capitata* and 15 *B. oleracea* var. *italica* populations, respectively, immersing microspore-derived haploid plantlet roots of *B. oleracea* var. *capitata* or *B. oleracea* var. *italica* in colchicine can artificially double chromosomes to over 50%. Formation of adventitious roots (AR) is of great importance for vegetative propagation, but difficult to achieve in many crop species. Aeroponic systems with varying root zone temperatures without using any plant hormones can successfully generate AR from stem-segment explants of *Brassica* species in less than a week. In *B. alboglabra*, *B. oleracea* var. acephala, *B. rapa* var. nipposinica, and *B. rapa* ssp. chinensis, maintaining cool root zone temperature (C-RZT; $20 \pm 2^{\circ}$C) and ambient root zone temperature (A-RZT; $30 \pm 2^{\circ}$C) significantly control the water and nutrient uptake capacity and stomatal conductance and differentially contribute to their productivity (Srikanth et al.).

STRESS MANAGEMENT AND MINERAL NUTRITION

Plant salt tolerance involves very complex mechanisms and varies with plant developmental and polyploidy levels. Notably, *Brassica* genotypes differing in ploidy level vary in their tolerance to salinity stress, where microsatellite (SSR) markers can help in their molecular breeding for salt tolerance (Kumar et al.). *Brassica* and *Arabidopsis* share the same family Brassicaceae, wherein an SSR marker-assisted comparative molecular marker mapping can contribute to the development of robust and healthy plants with high yield potential. *AtDjA3* gene encoding a heat shock protein 40 (J-protein) can be modulated by salt and osmotic stress. Notably, exhibition of differential seed morphology and sensitivity to salt, glucose, and ABA in *Atdja3*-null mutant line (j3) involves high transcript levels of ABA-insensitive 3 (Salas-Muñoz et al.). Among the ten classes of glutathione transferases (GSTs), known for their diversified important roles in plants, the phi, tau, lambda, and DHAR classes of GSTs are considered unique to plants. *Capsella rubella*, a member of the mustard family and a close relative of *A. thaliana*, possesses 49 GST genes coming within eight classes (He et al.). Additionally, *Capsella* and *Arabidopsis* GSTs exhibit functional divergence (both in gene expression and enzymatic properties) in paralogous gene pairs in *Capsella* (even the most recent duplicates), and orthologous GSTs in *Arabidopsis*/*Capsella*.

The nitrogen-efficiency capacity of plants such as *B. juncea* is modulated by N-supply and elevated [CO_2] conditions, where proteins like *PII*-like protein, cyclophilin, elongation factor-TU, oxygen-evolving enhancer protein, and rubisco activase are involved in maintaining photosynthesis, energy metabolism, and overall plant health (Yousuf et al.). Though regarded earlier as an environmental pollutant and a biotoxic agent, H_2S can play multiple functions in plants and improve cellular pools of nutrients including Ca, Cu, Fe, Mg, Mn, Na, and P under deficiency of certain element such as S (Reich et al.). For instance, in *Brassica pekinensis* (Lour.) Rupr. cv. Kasumi F1, H_2S partially prevented S-deficiency caused increase in the levels of Mo (and also Zn) that was argued a result of partial down-regulation of the sulfate transporters. Among important Ca^{2+} sensor proteins, calmodulins (CaMs) bind to certain TFs such as calmodulin-binding transcription activators (CAMTAs) and play important roles in various plant disease resistances and abiotic stress tolerances. Notably, *B. napus*, a tetraploid of the two progenitors *B. rapa* and *B. oleracea*, possess 18 CAMTAs, the highest number of CAMTAs among over 40 plant species studied for the same, and also 3-folds as many as that in *Arabidopsis* (Rahman et al.). *Arabidopsis* CAMTA mutant (AtCAMTA3) negatively regulates the resistance to *Sclerotinia sclerotiorum*, an ascomycete necrotrophic fungus, which induces the expression of *BnCAMTA3A1* and *BnCAMTA3C1*.

ENVIRONMENTAL PERSPECTIVES

In addition to its role in plant growth, the separated (monoculture) and combined (simultaneous) cropping of plants controls important processes at rhizosphere, solubility/bioavailability and accumulation/remediation of contaminants. Co-cropping of an *Alyssum murale* Ni-hyperaccumulator ecotype with nonhyperaccumulator *A. montanum* and perennial ryegrass (*Lolium perenne*) has no affect on Ni accumulation in *A. murale* leaves and stems (Broadhurst

and Chaney). However, co-cropping of *L. perenne* with *A. murale* can help the former to accumulate Cu up to 10 mg kg^{-1}, independent of the phytosiderophores. Additionally, Mn mobilization by the *Alyssum* hyperaccumulator species can significantly increase Mn levels in *L. perenne*. *Sinapis alba*, known as yellow/white mustard, is an important cruciferous condiment crop and also contributes to environmental pollutants-cleanup. In *S. alba* leaves, stems, and roots by *de novo* transcriptome analysis revealed the expression of 3,489, 1,361, and 8,482 unigenes, respectively (Zhang et al.). Genes pre-dominantly expressed in the leaves were enriched in photosynthesis- and C fixation-related pathways; whereas, stem-dominant genes were related with pathways related to sugar, ether lipid, and amino acid metabolisms and plant hormone signal transduction and circadian rhythm pathways. On the contrary, the root dominant genes were enriched in pathways related to lignin and cellulose syntheses involved in plant pathogen interactions, and potentially responsible for heavy metal chelating and detoxification. Though the oxidation of Glutathione (GSH) into GSSG occurs in leaves, its extensive conversion into PCs takes place in roots. Clues can be useful in the research and molecular breeding of *S. alba* and also for the molecular-assisted transfer of beneficial traits to other crops.

CONCLUSIONS AND OUTLOOK

Most of the contributions discussed molecular-genetic insights into agri-horticultural aspects that can help in devising breeding approaches in the members of the Brassicaceae family. Stress management and nutrition in *Brassica*, the second most discussed aspect, had little discussion on the environmental perspectives. Overall, this research topic yielded notable results/observations related to the yield and nutritional quality of and insights important for breeding *Brassica*. Studies on QTLs for branch angle, quantitative trait PDgr, and rRNA genes can significantly improve photosynthesis efficiency, biomass, flower and silique, and overall yield in *Brassica*. Studies are imperative for unveiling molecular insights into low N supply and elevated [CO_2] influences on the major plant proteins/enzymes and H_2S-mediated minimization of deficient-S accrued impacts in *Brassica*. The co-cropping of nonhyperaccumulator and metal-hyperaccumulator ecotypes has yielded promising results; however, more exhaustive studies are required in this direction to get benefit of *Brassica* in the environmental perspective. It is suggested, herein, to conduct more intensive research on breeding crops within the Brassicaceae family for improved health, productivity, and quality. In addition, more insights into the contribution of Brassicaceae family members to environmental issues also need to be addressed in future studies on the subject.

AUTHOR CONTRIBUTIONS

NA and SG prepared the first draft of the manuscript. OD, JJ, and NT read and revised the manuscript, and all authors listed approved the final version for publication.

ACKNOWLEDGMENTS

NA gratefully acknowledges the partial financial supports received from FCT (Government of Portugal) through contracts (SFRH/BPD/64690/2009 and SFRH/BPD/84671/2012), the Aveiro University Research Institute/CESAM (UID/AMB/50017/2013), and COMPETE through Project no. FCOMP-01-0124-FEDER-02800 (FCT PTDC/AGR-PRO/4091/2012). SG also acknowledges the partial financial supports received from University Grants Commission (UGC) and Council of Scientific and Industrial Research (CSIR), Govt. of India, New Delhi.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Anjum, Gill, Dhankher, Jimenez and Tuteja. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 September 2016
doi: 10.3389/fpls.2016.01369

Floral Initiation in Response to Planting Date Reveals the Key Role of Floral Meristem Differentiation Prior to Budding in Canola (*Brassica napus* L.)

Yaofeng Zhang, Dongqing Zhang, Huasheng Yu, Baogang Lin, Ying Fu and Shuijin Hua**

Institute of Crop and Nuclear Technology Utilization, Zhejiang Academy of Agricultural Sciences, Hangzhou, China

In *Brassica napus*, floral development is a decisive factor in silique formation, and it is influenced by many cultivation practices including planting date. However, the effect of planting date on floral initiation in canola is poorly understood at present. A field experiment was conducted using a split plot design, in which three planting dates (early, 15 September, middle, 1 October, and late, 15 October) served as main plot and five varieties differing in maturity (1358, J22, Zhongshuang 11, Zheshuang 8, and Zheyou 50) employed as subplot. The purpose of this study was to shed light on the process of floral meristem (FM) differentiation, the influence of planting date on growth period (GP) and floral initiation, and silique formation. The main stages of FM developments can be divided into four stages: first, the transition from shoot apical meristem to FM; second, flower initiation; third, gynoecium and androecium differentiation; and fourth, bud formation. Our results showed that all genotypes had increased GPs from sowing to FM differentiation as planting date was delayed while the GPs from FM differentiation to budding varied year by year except the very early variety, 1358. Based on the number of flowers present at the different reproductive stages, the flowers produced from FM differentiation to budding closely approximated the final silique even though the FM differentiated continuously after budding and peaked generally at the middle flowering stage. The ratio of siliques to maximum flower number ranged from 48 to 80%. These results suggest that (1) the period from FM differentiation to budding is vital for effective flower and silique formation although there was no significant correlation between the length of the period and effective flowers and siliques, and (2) the increased number of flowers from budding were generally ineffective. Therefore, maximizing flower numbers prior to budding will improve silique numbers, and reducing FM degeneration should also increase final silique formation. From the results of our study, we offer guidelines for planting canola varieties that differ in maturity in order to maximize effective flower numbers.

Keywords: canola yield, floral meristem, floral induction, planting date, floral organs

OPEN ACCESS

Edited by:
Sarvajeet Singh Gill,
Maharshi Dayanand University, India

Reviewed by:
Maoteng Li,
Huazhong University of Science and Technology, China
Kamrun Nahar,
Kagawa University, Japan

****Correspondence:***
Shuijin Hua
sjhua1@163.com
Dongqing Zhang
dongqing_zhang@126.com

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *16 March 2016*
Accepted: *29 August 2016*
Published: *14 September 2016*

Citation:
Zhang Y, Zhang D, Yu H, Lin B, Fu Y and Hua S (2016) Floral Initiation in Response to Planting Date Reveals the Key Role of Floral Meristem Differentiation Prior to Budding in Canola (Brassica napus L.). Front. Plant Sci. 7:1369. doi: 10.3389/fpls.2016.01369

INTRODUCTION

Planting date is a simple but essential agronomic practice during crop production. Recommendation of optimal planting date depends on the combination of several factors including plant variety, temperature suitability, and water availability. In modern canola (*Brassica napus* L.) production systems, planting date should be re-considered because of climatic changes, newly bred canola varieties, modern agricultural developments, and human social activities. Considering climatic changes, global warming is the most notable (Manciocco et al., 2014; Yang and Chen, 2014). One of the consequences of global warming is alteration of the growth period (GP) in crops such as rice (Huang et al., 2009). Because the rice-canola rotation is a widely practiced cropping system in China (Li et al., 2012), delaying the harvest time for rice can correspondingly affect the planting date for canola. To the best of our knowledge, newer canola varieties should be adapted to certain ecological areas, which can vary in terms of temperature, amount of sunshine, moisture availability, etc. Hence, the goal in breeding new varieties must generally meet those requirements. Once a new canola line has been developed, evaluation of many traits, including planting date, is necessary to obtain optimal yield and quality before it is released. As for modern agricultural developments and human social activities, important progress in canola production in China has been achieved by the introduction of mechanization for sowing, fertilizing, and harvesting. Because mechanization is labor saving and highly efficient compared to manual practices (Lu, 2009; Hormozi et al., 2012), many agronomic practices are being replaced, such as substitution of transplanting by direct seeding. However, planting dates for canola differ considerably between the two methods because of differences in the seedling GP from sowing to transplanting. Furthermore, the physiological status of the seedlings is also different between the two sowing practices (Boyhan et al., 2009; Mulyati et al., 2009). Thus, it is essential to re-estimate the planting date of canola varieties from the above-analyzed cases.

Normally, either early or late planting dates are adverse for obtaining high yields. Early planting usually results in larger plants. However, there is a trade-off in the consumption of more fertilizer and moisture (Gormus and Yucel, 2002; Sorizano et al., 2004). It has been reported that delayed planting can reduce yield in many crops, including canola (Chen et al., 2005; Sindelar et al., 2010; Tsimba et al., 2013). In our recent study, we found that the reduction of seed yield in canola was largely associated with a decrease in the number of siliques (Hua et al., 2014). The *B. napus* silique is derived from floral meristem (FM) differentiation, and thereafter by sequential flower organ development and normal fertilization (Martínez-Laborda and Vera, 2009; Ferrándiz et al., 2010). In this context, factors that affect FM differentiation should be directly correlated with the final number of siliques. Plant FM differentiation is determined by a population of stem cells within the shoot apical meristem (SAM; Souer et al., 1996). The SAM is involved in organogenesis, and produces aboveground plant tissues such as new leaves and axillary buds (Bowman and Eshed, 2000). Thereafter, the SAM is ultimately responsible for plant architecture. There are two main processes that occur in the SAM: (1) transition from the SAM to the FM, and (2) FM differentiation. These processes have been well characterized in *Arabidopsis* (Irish, 2010). Differentiation of stem cells in the SAM is regulated by many factors including developmental cues, phytohormones, the environment (i.e., temperature and day length), and their interactions (Pickett et al., 1996; Jouannic et al., 2011; Takacs et al., 2012; Uchida et al., 2013). Although several key factors have been identified that function in stem cell transition and differentiation of the SAM and FM, i.e., *LEAFY*, *APETALA1*, *AGAMOUS*, and *WUSCHEL* (Williams et al., 2005; Yang et al., 2013), knowledge on how the plant perceives environmental cues such as temperature and nutrient supply, and responses to crop nursery practices, i.e., planting date and density, are largely unknown.

When modifying planting date, the major issue encountered for canola genotypes is to readjust its physiological status based on changes in climatic factors such as temperature and sunshine hours. For example, the mean temperature in September is much higher than it is in October in China (Hua et al., 2014). High temperatures might tend to modulate physiological metabolism in the plant cell (Suwa et al., 2010; Djanaguiraman et al., 2014). However, will the altered processes that result from different planting dates affect the timing of the transition from SAM to FM and FM differentiation? If so, then will the variation in FM differentiation affect silique formation? Therefore, the profound impacts of changes in planting date on floral initiation should be investigated, because they are closely associated with canola yield. Hence, the goals of the present study were to (1) dissect the morphology of FM during differentiation, (2) compare the timing of the transition of the SAM to FM and FM duration with respect to different planting dates, (3) analyze the variation in the number of flowers at different growth stages with respect to different planting dates, and (4) define the relationship between FM differentiation and effective silique number.

MATERIALS AND METHODS

Site Description and Crop Management

The experiment was carried out during the 2011–2012 and 2012–2013 growing seasons at the experimental station of the Zhejiang Academy of Agricultural Sciences, Hangzhou, China. Five canola (*B. napus* L.) varieties with different maturities; 1358 (very early), J22 (early), Zhonshuang 11 (middle), Zheshuang 8 (late), and Zheyou 50 (late), were chosen as plant materials. Seeds of 1358, Zhongshuang 11, and Zheshuang 8 were provided by Professor Chunyun Guan, Hunan Agricultural University, while J22 and Zheyou 50 were bred by our research group. The soil type in the experimental station is loamy clay (loamy, mixed, and thermic Aeric Endoaquepts). Canola was rotated with rice and thus the crop previously cultivated was rice. Before sowing, fertilizers including urea, calcium superphosphate, potassium oxide, and borax were broadcast at the rate of 275, 375, 120, and 15 kg ha^{-1}, respectively, as a basal fertilizer dose. In addition, urea at the rate of 120 kg ha^{-1} was applied as a topdressing at the end of January in 2012 and 2013. Approximately five canola seeds

were directly sown into the soil in each shallow hole in the plot at a depth of approximately 3 cm. After 1 month, the seedlings were thinned to one plant hole^{-1}. The field was not irrigated during the canola growing season, and the rainfall profile for the two growing seasons in shown **Figure 1**. Weeds were manually removed at the seedling stage, and aphids were controlled by application of omethoate emulsion 0.06% (V/V) when the plants were finished flowering.

Experimental Design

The experiment was a split plot arrangement in a randomized complete block design with three replications. Three planting dates; early (15 September), optimal (1 October), and late (15 October), were employed as the main plot while five canola genotypes with different maturities; very early (1358), early (J22), middle (Zhongshuang 11), and late (Zheshuang 8 and Zheyou 50), were surveyed as sub plot. The recommended optimal planting date was considered as the standard with which to compare the effects of different planting dates on agronomic traits and floral initiation. The plants were grown in plots (40 m in length), with spacing between the rows of 0.35 m and spacing between the plants at 0.2 m.

Imaging the FM Differentiation Process

To record the FM differentiation process, 10 plants were taken from each plot and delivered to the laboratory for observation of FM differentiation. Roots of the sampled plants were rinsed with water to remove the soil, and all leaves were then removed until they were difficult to distinguish with the naked eye. The removal of the developmental leaves was then performed with a stereomicroscope (Leica M205 C, Germany) using sharp forceps. The growing tip was observed under magnification, and all leaves that enclose the FM were then removed. The morphology of the SAM, FM, and the process of FM differentiation were photographed using software that accompanied the stereomicroscope. The criterion for establishment of a specific differentiation stage (interpreted in the Results section, **Figure 2**) was that $>50\%$ of plants enter the given stage.

Determination of Plant Growth Periods (GPs) for Canola

Data for the canola plant GP was recorded for the different planting dates and included sowing, transition from SAM to FM, budding (BBCH51, Lancashire et al., 1991), initial flowering (BBCH60), middle flowering (BBCH65), end of flowering (BBCH69), maturation (BBCH89), and harvesting (BBCH99). The GP was defined as the interval between two sequential stages. The estimation standard for each growth stage is described as follows: (1) transition from SAM to FM – a first lateral outgrowth was observed on the periphery of meristem; (2) budding – a flower bud >1 cm was observed; (3) initial flowering – opening of the first flower on a plant; (4) middle flowering – approximately 50% percent of the flowers were open on a plant; (5) end of flowering – all the flower buds were open and the plant had finished flowering. At each stage, when 50% of the plants reached the criterion as described above, it was considered to have entered the designated stage.

Measurement of Flower Numbers

Canola plants were sampled at the budding, initial flowering, middle flowering, end of flowering, and harvest stages. Five plants in each plot were sampled. The flower buds were counted in a combination of those readily visible to the naked eye and those observed through a stereomicroscope and recorded at each sampled stage. When identified through the stereomicroscope, flower primordia were also recorded as flowers (**Figure 2c**). After cessation of flowering, when siliques were formed, all effective and ineffective siliques and unopened flowers were included during counting. For ease of description, we use the terminology 'flower numbers' to include siliques, flowers, and unopened flowers. To reduce the large amount of work required, flowers and siliques were counted on the main inflorescence and the 1st, 4th, and 8th branches on the canola plants, and this included branches from different stem positions.

Statistics

In this experiment, analysis of variance (ANOVA) on flower numbers, silique numbers, and the ratio of siliques to the maximum number of flowers was analyzed using the SAS PROC MIXED procedure (SAS Institute, 2004), where planting dates and canola genotypes were considered to be fixed effects and the year (because the experiments were conducted in the same plots), and replication and all their interactions were considered to be random effects. Further comparisons of multiple means for significance on flower numbers under three planting dates and five genotypes (Supplementary Table S1) at each developmental stage including budding, initial flowering, middle flowering, end of flowering, and maturation were performed using Duncan's method (P-value < 0.05). Furthermore, siliques at harvest stage, and the ratio of siliques to the maximum number of flowers of five genotypes on three planting dates were compared using Duncan's method (P-value < 0.05) as well.

RESULTS

Mean Temperature, Precipitation, and Sunshine Hours during the Canola Growing Seasons

Growth of canola requires optimal temperatures, soil moisture (precipitation), and light. However, these climatic factors constantly fluctuate, and their interactions differ from year to year. Thus, the canola plants need to adjust their responses to cope with the environmental changes. Mean temperature varied between the two growth seasons (**Figure 1A**). The apparent differences occurred during four time periods; in November, from 20 December to 10 January, from 20 January to 20 March, and from 20 March to 20 May. The four time periods during 2011–2012 showed higher mean temperature except from 20 January to 20 March. During the temperature recovery stage, which is normally from the end of January, although both years

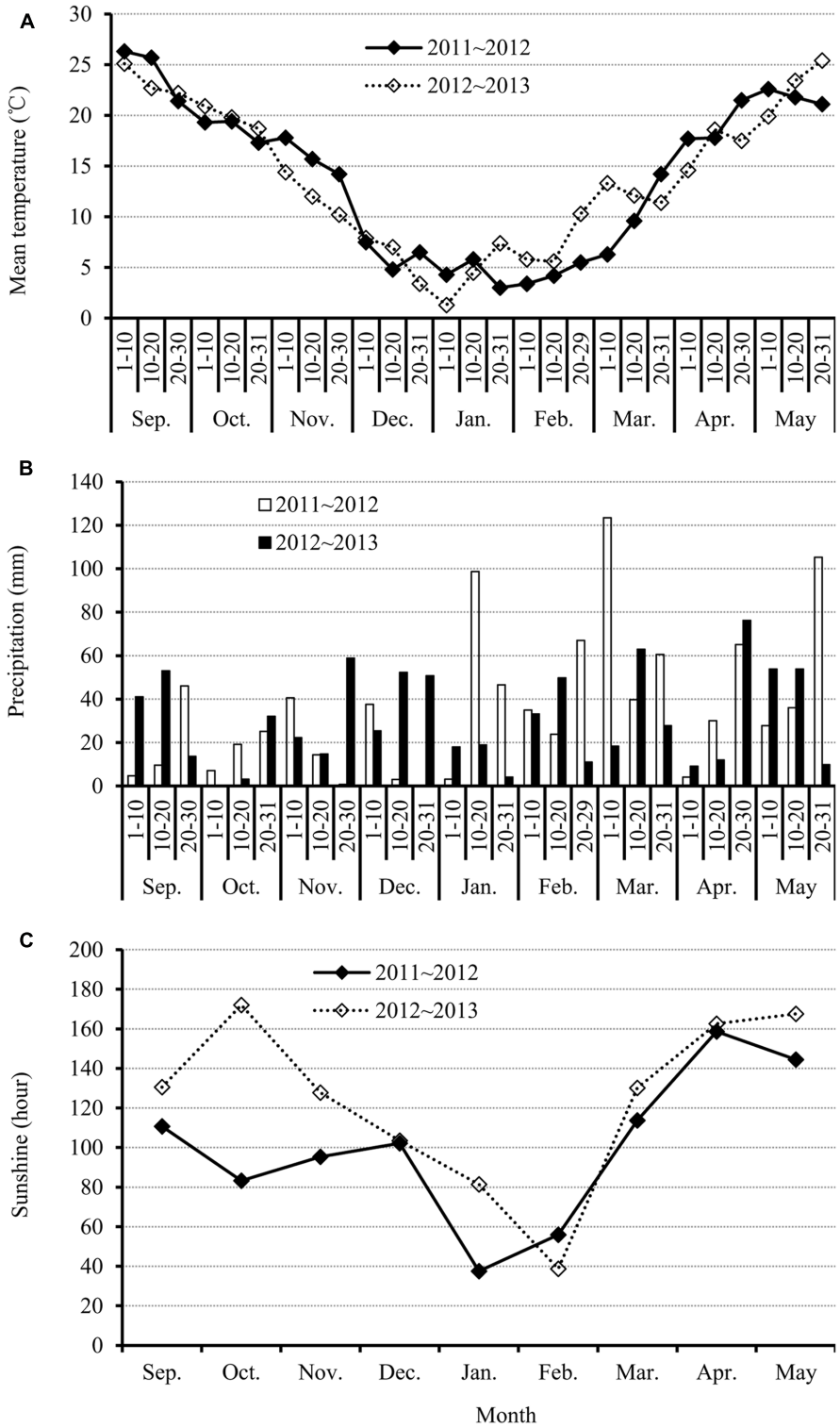


FIGURE 1 | Dynamics of climatic factors including (A) mean temperature, (B) precipitation, and (C) sunshine hours for each month during the 2011–2012 and 2012–2013 canola growing seasons.

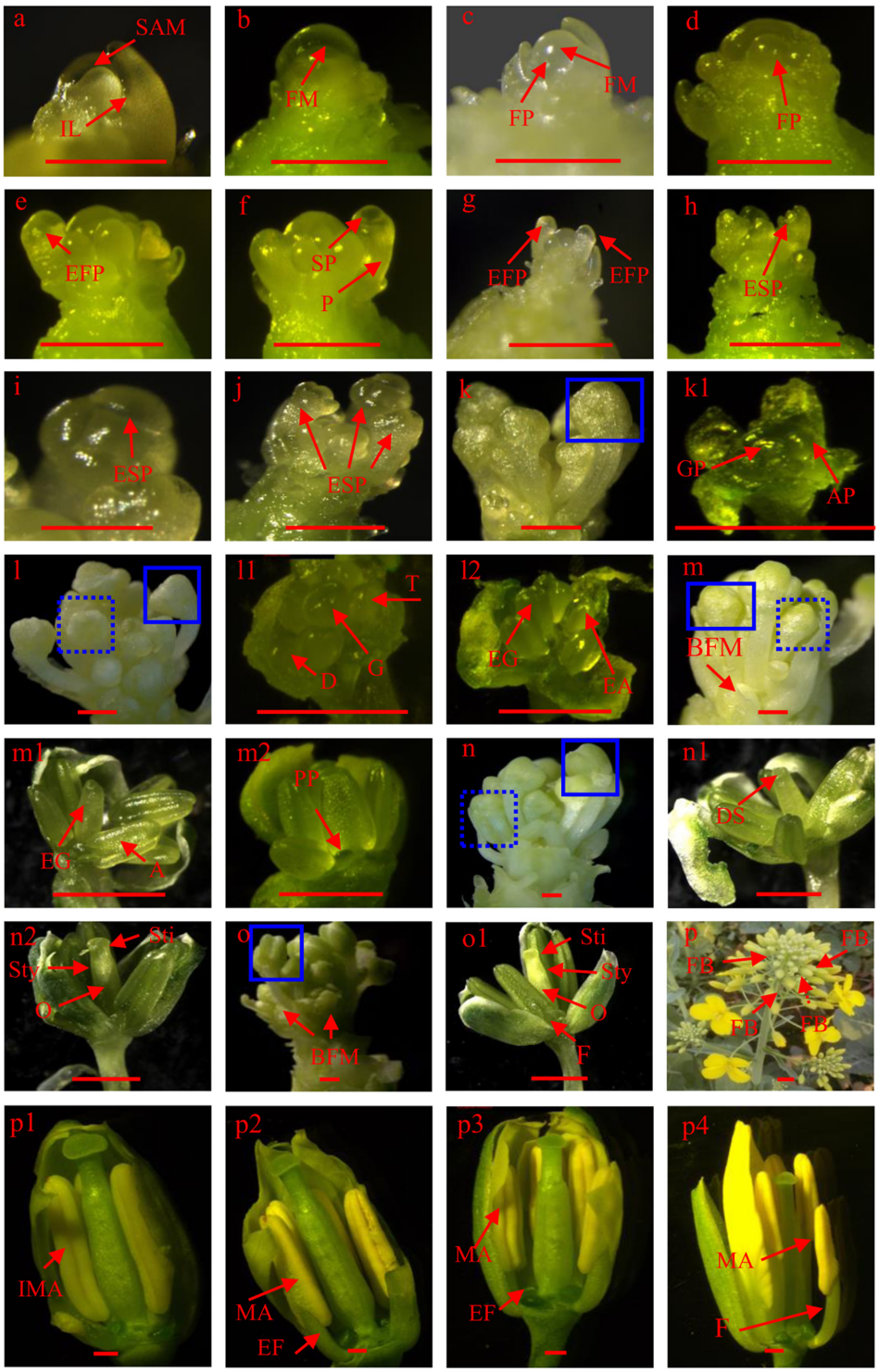


FIGURE 2 | Floral meristem (FM) differentiation: (a) shoot apical meristem, (b) FM, (c) and (d) floral primordium production on the periphery, (e–j) sepal primordium differentiation, (k–l2) gynoecium and androecium primordium differentiation, (m–o1) petal primordium differentiation and gynoecium and androecium further development, (p) flower bud formation, (p1–p4) anther and stigma maturation. A, anther; AP, androecium primordium; BFM, branch floral meristem; D, didynamous stamen; DS, dented stigma; EA, elongated androecium; EF, elongated filament; EFP, elongated floral primordium; EG, elongated gynoecium; ESP, elongated sepal primordium; F, filament; FB, flower bud; FM, floral meristem; FP, floral primordium; G, gynoecium; GP, gynoecium primordium; IL, innermost leaf; IMA, immature anther; MA, matured anther; O, ovary; P, pedicel; PP, petal primordium; SAM, shoot apical meristem; SP, sepal primordium; Sti, stigma; Sty, style; T, tetradynamous stamen. Bars equal 1 mm.

showed increasing trend, the mean temperatures in 2012–2013 were much higher than in 2011–2012. The maximum difference of mean temperature between the 2 years reached 7°C. The fluctuations in temperature during the canola growing season could affect many developmental processes, including floral organ initiation.

For precipitation in the two growth seasons, there were several properties observed in **Figure 1B**. First, the precipitation was distributed very unevenly in different months over the year. For example, the amount of precipitation in January and March was considerably more than in other months in 2011–2012. Second, the differences in precipitation in the same month between the two growing seasons were very large. For example, in certain periods, such as from 20 to 30 November, precipitation in 2012–2013 reached ∼60 mm while in 2011–2012 it was zero. However, some periods showed the opposite trend. For instance, a very high level of precipitation was measured (∼100 mm) from 10 to 20 January in 2011–2012, but it was only ∼20 mm in 2012–2013 (**Figure 1B**).

Sunshine hours exhibited a decreasing trend from autumn to winter, but then increased, from winter to spring, in both growth seasons (**Figure 1C**). However, the number of sunshine hours from September to November in 2011–2012 were much less than in 2012–2013, especially in October, which showed an increase of >80 h of sunshine in 2012–2013 (**Figure 1C**).

The data in **Figure 1** clearly shows that climatic factors, mainly the mean temperature, precipitation, and sunshine hours, were considerably different in the two growing seasons and the canola plants need to adjust their responses to this environmental variation.

Floral Meristem (FM) Differentiation in Canola

The major steps in canola FM differentiation are illustrated in **Figure 2** which can be divided four main stages. First, SAM transits to FM. Pior to differentiation, the SAM appears conical in shape (**Figure 2a**). Morphological landmarks in the transition from SAM to FM are (1) the meristem becomes much rounder, and (2) bulges form on the periphery of the meristem (**Figures 2b,c**). Second, flower initials from flower primordium. Once FM differentiation has initiated, many outgrowths, namely flower primordia, are produced and surround the meristem (**Figure 2d**). As more new primordia are produced, the outermost flower primordium begins to elongate, and differentiates into a pedicel from the basal part to the bubbled flower primordium (**Figure 2e**). After the pedicel has further elongated, a new protrusion forms in the middle of the bulge, which is the sepal primordium and will develop into a sepal (**Figure 2f**). As elongation of the flower primordium increases (**Figure 2g**), the sepal primordium gradually elongates as well (**Figure 2h**). The four sepals do not develop at the same speed because two of them are obviously longer than the others (**Figure 2i**). Third, gynoecium and androecium differentiate. After the bulge is thoroughly enclosed by sepals (**Figure 2k**), the gynoecium and androecium rapidly undergo differentiation (**Figures 2k,k1**). However, we observed that there are only four androecia. Furthermore, the morphology of the gynoecium and androecium primordia was very similar and was bulged as well (**Figures 2k,k1**). As development of the flower bud progressed, the longer sepals covered the bulged flower primordium (**Figure 2l**), didynamous stamens formed in the younger flowers, and the top of the gynoecium became dented (**Figure 2l1** and the dashed box in **Figure 2l**). In older flowers, both the gynoecium and androecium elongated and a vertical dent appeared in the androecium (**Figure 2l2** and solid box in **Figure 2l**). As the flower bud aged, the shorter sepals became further elongated and branches also began to develop on the FM (**Figure 2m**). At this stage, the gynoecium was slightly taller than the androecium and the top of the gynoecium was still dented in younger flowers (**Figure 2ml** and the dashed box in **Figure 2m**). However, the lengths of the gynoecium and androecium are almost the same in older flowers. Moreover, the petal primordium was easily observed. Until this time, the elongated genoecium was cylindrical (**Figure 2m2** and the solid box in **Figure 2m**). Characteristic of the more developed flower bud was that the four sepals were very apparent and the lengths became similar (**Figure 2n**). The upper part of the gynoecium developed a crack and was much thinner than the lower part in younger buds (**Figure 2n1** and the dashed box in **Figure 2n**). Three parts, the stigma, style, and ovary, were clearly divided into older flowers (**Figure 2n2** and the solid box in **Figure 2n**). Before budding, older flowers become much more mature (**Figure 2o1** and the solid box in **Figure 2o**). Furthermore, the stigma, style, and ovary are more well defined, and the filaments start to differentiate in older flowers (**Figure 2o1**). At this point, all organs in the older buds of the main inflorescence are fully formed and awaiting suitable temperatures to induce budding. Fourth, bud forms. After budding (**Figure 2p**), the main variations in the flower organs are found in the androecium. As the buds age, the anthers are initially yellow–green in color and the filaments are short (**Figure 2p1**). The anthers then turn yellow. Although the filaments become longer, the length of the anther and the filament is still shorter than the stigma at the p2 stage. At the p3 stage, the anther and stigma are almost the same length. Before the flowers open, the mature anther is above the stigma, which ensures pollen grain release and pollination (**Figure 2p4**).

Influence of Planting Date and Genotype on Growth Periods (GPs) in Canola

The GP represents an important developmental stage for plant organ development, environmental response, and nutrient cycling.

Generally, the total growth periods (TGPs) for canola decreased as the planting date were delayed in all genotypes (**Figure 3**). The TGP is composed of six stages; seed sowing to FM differentiation (s1), FM differentiation to budding (s2), budding to initial flowering (s3), initial flowering to middle flowering (s4), middle flowering to the end of flowering (s5), and end of flowering to harvest (s6).

The s1 GP in canola varieties increased with the increase in plant maturity for each respective planting date. This result

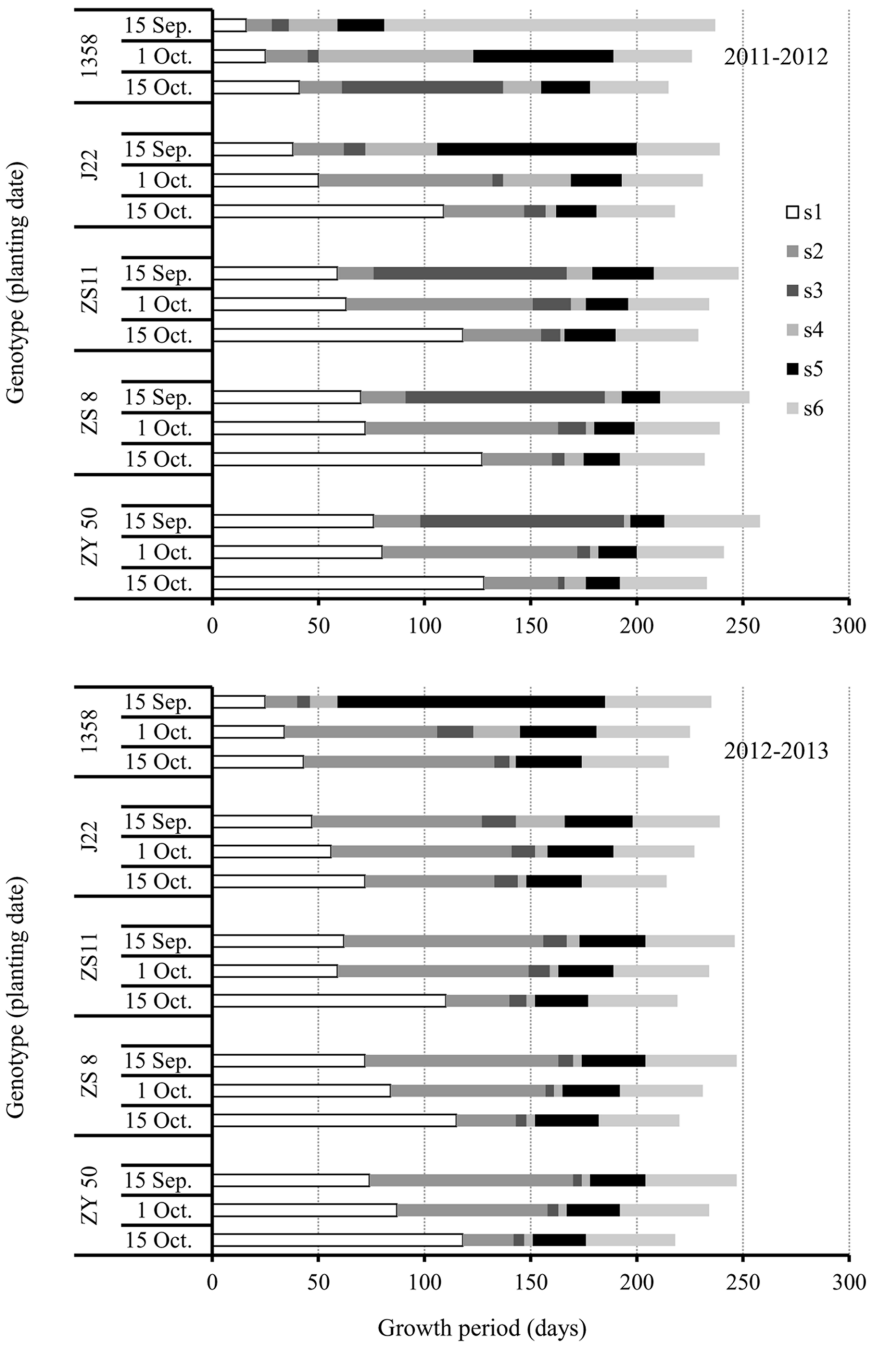


FIGURE 3 | Growth periods (GPs) during growth of the five canola varieties, 1358, J22, Zhongshuang 11 (ZS11), Zheshuang 8 (ZS8), and Zheyou 50 (ZY50) for three planting dates; early (September 15), optimal (October 1), and late (October 15) in the 2011–2012 (top) and 2012–2013 (bottom) growing seasons. s1, GP from sowing to initiation of FM differentiation; s2, GP from initiation of FM differentiation to budding; s3, GP from budding to initiation of flowering; s4, GP from initiation of flowering to middle flowering; s5, GP from middle flowering to the end of flowering; s6, GP from the end of flowering to harvest.

suggested that early maturing canola varieties, such as 1358, completed the transition process from SAM to FM or from vegetative to reproductive growth much quicker than late maturing varieties, and had no strict requirement for low temperature. As for the impact of planting date on the transition from SAM to FM, our results showed that the s1 GPs increased as planting date was delayed in all varieties. However, the extension of the s1 GPs for the late planting date was especially significant. The result suggested that low temperature under late planting can delay the transition from SAM to FM.

The s2 GPs (from FM to budding) varied greatly among years, genotypes, and planting dates (**Figure 3**). In 2011–2012, varieties showed the longest s2 GPs at recommended plant date except 1358, which had the similar s2 GPs for both the optimal and late planting dates. However, the response of s2 GPs to planting date in all canola varieties was very different in the 2012–2013 growing season. For 1358, the s2 GPs increased significantly as the planting date was delayed. Another early maturing variety, J22, showed the longest s2 GP for the recommended planting date and the shortest s2 for the late planting date. However, in ZS11, ZS8, and ZY50, the lengths of the s2 GPs decreased as the planting dates were delayed in 2012–2013. The s2 GPs for the early planting dates were three–fourfold longer than for the late planting dates in ZS11, ZS8, and ZY50. The result revealed that the period from FM to budding was heavily depending on the climatic factors.

The GPs from s3 to s6, J22, ZS11, ZS8, and ZY50 showed much shorter than that of s1 and s2 at optimal and late planting date in 2011–2012 while they showed the same behavior in all planting dates in 2012–2013. For the early maturing variety 1358, the adjustment of GPs was considerable in various years. For example, the GP from end of flowering to maturation (s6) in the 2011–2012 growth season was very long for this variety, while the longest period was from middle flowering to end of flowering (s5) in the 2012–2013 growth season.

Taken together, although TGPs were similar between the two growing seasons in canola varieties that differed in maturity, the individual GPs were greatly influenced by planting date and year.

Impact of Planting Date and Genotype on Flower Numbers

Flower numbers were counted from budding to maturation stages in the five canola varieties. Result of ANOVA showed that effect of planting date on flower numbers were not significant but significantly affected by genotype, year and their interactions at budding stage (**Table 1**). At other stages, flower numbers were strongly influenced by planting date and genotype and their interactions (**Table 1**).

Results showed that flower numbers increased as the canola plants developed, and that flower numbers generally peaked at the end of the flowering stage (**Figures 4A,B**). This showed that FM differentiation is a continuous process that does not end at the budding stage. Considering the effect of planting date on flower numbers, we found the fewest flowers on plants from the late planting date except for the very early variety 1358 (**Figures 4A,B**). In fact, plants of variety 1358 from the early planting date had the fewest flowers. This result suggests that the adverse impact of early planting on very early maturing varieties is much more profound than it is for later planting dates, while the opposite trend was observed for the other four varieties.

In examining the influence of genotype on flower numbers, we found that ZS8 had the maximum number of flowers for the early planting date, while 1358 had the fewest flowers in the 2011–2012 growing season (**Figures 4A,B**). However, the early variety J22 had the maximum number of flowers for the early planting date while ZY50 had the fewest in the 2012–2013 growing season. The relative differences between the 2 years were 100 and 50 flowers, respectively. However, when the flowers were counted at harvest, we found that ZS8 and ZY50 ranked first and second in both years for the recommended planting date (**Figures 4A,B**; Supplementary Table S1).

The results of our study show that the appropriate planting date for very early maturing canola varieties is 1 October because both early and late planting dates can reduce the number of flowers. However, the early-, middle-, and late-flowering canola varieties had the greatest potential to produce more flowers at early planting date.

Effect of Genotype and Planting Date on Silique Formation

The eventual outcome of flower primordium differentiation is silique formation. Therefore, we counted the total number of siliques on the main inflorescence and the first, fourth, and eighth branches at the harvesting stage.

In general, more siliques were formed on plants from the recommended planting date than from any other planting date (**Figure 5**). However, the effect of early and late planting dates on silique numbers depended on the canola genotype. For example, the very early variety, 1358, had more siliques from the late planting date, and showed a 10% increase on average for the 2 years as compared with the early planting date. For the other canola varieties, the early planting date was advantageous for silique formation in comparison with the late planting date. Considering genotypic variation, ZS8 had the most siliques, producing 264 siliques plant^{-1} on average over the 2 years, with a total of four inflorescences, for the recommended planting date. The very early maturing variety produced the fewest siliques, averaging 200 siliques over the 2 years for the recommended planting date. Therefore, both planting date and genotype can affect silique production in *B. napus*.

The ratio of siliques to the maximum number of flowers was also recorded. The results showed that regardless of the similarity in the number of siliques formed on plants from different planting dates with different genotypes, the ratio of siliques to the maximum number of flowers varied drastically between the 2 years (**Figure 5**). In general, all canola varieties, except for the very early maturing variety (1358), exhibited higher silique formation ratios for the late planting date in 2011–2012. However, the ratio of siliques to the maximum number of flowers for the late planting date in 2012–2013 was less than for the other two planting date for ZS11, ZS8, and ZY50. The higher silique formation ratio in the 2011–2012 growing season was due to the lower maximum silique numbers. J22 had a higher silique formation ratio for the late planting date, exceeding 80%, for both years. In addition, ZS8 had a higher silique formation ratio at the recommended planting date that averaged >80% in 2012–2013.

DISCUSSION

Our study focused on the morphology of FM differentiation, the timing and duration of FM differentiation, and floral organ initiation with respect to three different planting dates in canola.

TABLE 1 | Analysis for variances (planting date, genotype, and year) at different growth stages (budding, BBCH 51; initial flowering, BBCH60; middle flowering, BBCH65; end of flowering, BBCH69; and maturation, BBCH89) (*P* > *F*).

Source	Budding	Initial flowering	Middle flowering	End of flowering	Maturation
Planting date (PD)	0.488	<0.001	<0.001	<0.001	<0.001
Genotype (G)	<0.001	<0.001	<0.001	<0.001	<0.001
Year (Y)	<0.001	<0.001	0.292	<0.001	0.237
PD*G	<0.001	<0.001	<0.001	<0.001	<0.001
PD*Y	<0.001	0.074	0.268	<0.001	0.083
G*Y	0.018	<0.001	<0.001	<0.001	<0.001
PD*G*Y	<0.001	<0.001	<0.001	<0.001	<0.001

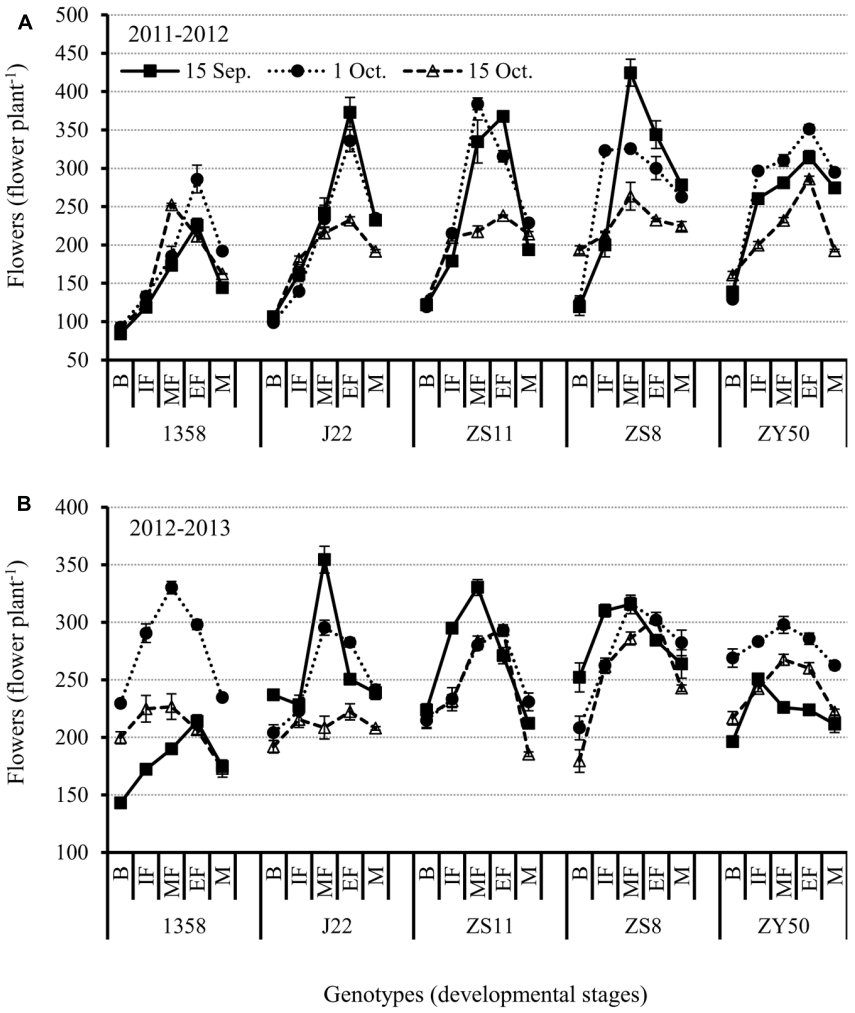


FIGURE 4 | Changes in flower numbers at different reproductive stages in the five canola varieties 1358, J22, Zhongshuang 11 (ZS11), Zheshuang 8 (ZS8), and Zheyou 50 (ZY50) for three planting dates; early (September 15), optimal (October 1), and late (October 15) in the 2011-2012 (A) and 2012-2013 (B) growing seasons. B, budding stage; IF, initiation of flowering stage; MF, middle flowering stage; EF, end of flowering stage; and M, maturation stage. Bars indicate standard errors of the mean.

Flower and silique formation are the most important issues for canola production because they directly affect seed yield. The adverse influence of delayed planting on canola seed yield has been assessed previously (Holman et al., 2011); however, there are fewer relevant reports describing how floral initiation is affected by different planting dates.

The flower is derived from a flower primordium. Before flower primordium differentiation, the meristem located in the

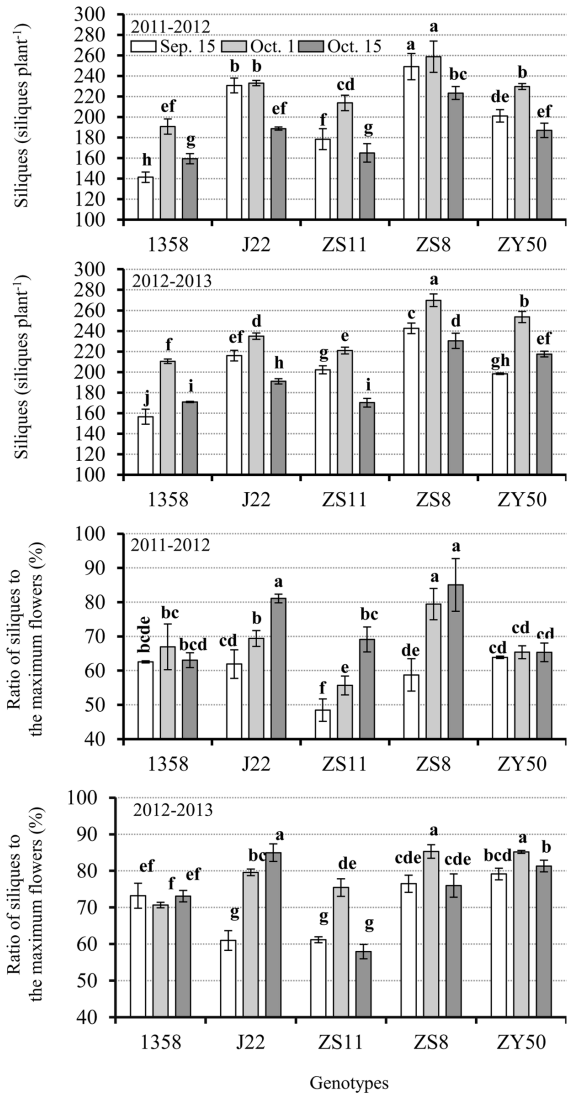


FIGURE 5 | Silique numbers and the ratios of siliques to the maximum number of flowers at harvest in the five canola varieties 1358, J22, Zhongshuang 11 (ZS11), Zheshuang 8 (ZS8), and Zheyou 50 (ZY50) for three planting dates; early (September 15), optimal (October 1), and late (October 15) in 2011–2012 and 2012–2013. Different letters on each histogram indicate significant differences at 5% level using Duncan's test.

inmost leaves is the SAM, not the FM (Irish, 2010). Therefore, only the SAM transitions into the FM, and a number of flower primordia can be produced and the floral organs such as sepal, pistil, anther, carpel, stigma, and petal will form sequentially. The morphology and process of FM differentiation among the different canola genotypes in this study are the same. Therefore, we only provide an overview of the process of FM differentiation. This suggests that the difference of canola genotypes normally occurs at the amount of flowers but not morphology. Initiation of the transition from SAM to FM is evident from **Figures 2a–c** as described in the section "Results." In *Arabidopsis*, similar morphological changes, such as an outgrowth produced on the flank of flower primordium, were observed (Smyth et al., 1990). Once the FM differentiation had initiated, the main processes observed were the production of new flower primordia on the periphery of the FM, and the development of the flower primordia into flowers through differentiation of each organ on the outermost flower primordia simultaneously. We observed that these processes in canola were generally similar to those in the related model plant *Arabidopsis thaliana* (Smyth et al., 1990). The other important issue was that of branch inflorescence formation during floral organ initiation in the main inflorescence (**Figures 2m,o**). Although we yet do not know how many branch inflorescences are formed during development of the main inflorescence, we deduced that this aspect of the branch inflorescence is genetically controlled. Furthermore, we also inferred that the branch inflorescences that developed at this stage were effective, because the lower position leaves also contain axillary buds, but they are normally dormant and cannot developed into branches. Branching is also an intricate regulatory network that is affected by several phytohormones such as auxin and cytokinin (Bainbridge et al., 2005; Stirnberg et al., 2012). Therefore, further analysis should focus on the formation of effective inflorescences during development of the main inflorescence to determine whether they are affected by agronomic practices such as planting date, because an increase in the number of effective branch inflorescences can also increase the total numbers of flowers and siliques.

The development of floral organs is affected by many environmental and agronomic practices. The planting date triggers the appropriate initiation and timing of FM differentiation. In canola, the transition from SAM to FM is the prelude to the reproductive stage. In our study, the very early variety showed rapid transition from SAM to FM. The quick transition from vegetative to reproductive growth can have severe consequences. First, rapid transition often results in lower accumulation of plant biomass and hence the nutrient supply from leaf and stem compared to the recommended planting date (unpublished date). During floral development, a considerable level of nutrients should be absorbed from other tissues (Sklensky and Davies, 2011). Consequently, flower organ formation may be predictably decreased. A second consequence is low temperature stress on flower and silique development. As seen in **Figures 1** and **3,** the very early maturing variety encountered low temperatures during flowering and silique development. The influence of low temperatures on flowering and silique development can be destructive (Wingler, 2011). We observed that developing seeds aborted after low temperature stress (**Figure 6**). Therefore, early planting should be avoided for the very early maturing canola varieties. Even though more flowers and siliques formed on plants from the late planting date than the early planting date for the very early maturing variety, late planting is not recommended. When the very early maturing variety was planted late, the timing from vegetative to reproductive stage increased because of decreasing temperatures in November. It has been reported that low temperature can retard or even terminate FM differentiation (Sun B. et al., 2009; Liang et al., 2012). However, although the low temperatures

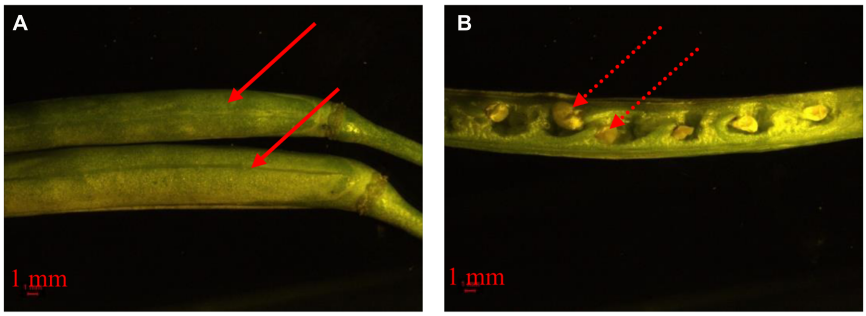


FIGURE 6 | Morphology of developing siliques (A) and seeds (B) in response to low temperature stress. The solid arrow means the siliques of the very early maturing variety, 1358, from different branches in early January. The dashed arrow means aborted developing seeds.

slowed the growth of the plants, the very early maturing variety is able to accelerate the initiation of FM differentiation when canola has been exposed to low temperatures to pass vernalization (Adhikari et al., 2012; Wollenberg and Amasino, 2012). Once a suitable temperature is reached, the canola plant can promptly differentiate many flowers with a small vegetative biomass as well. Unlike early planting, the vegetative growth is also vigorous for the very early maturing variety prior to the arrival of low temperatures; however, the very early variety may not develop much biomass, with fewer leaves and branches, which could then lead to reduced flower numbers as compared with plants grown from the recommended planting date.

For the other four canola varieties in this study, the timing of the transition from vegetative to reproductive growth in response to the late planting date was also longer than for either the early or recommended planting dates. At the early and recommended planting dates, the canola varieties required a period of low temperature for vernalization (Filek and Dubert, 1994; Sheldon et al., 2009). Thus, a low temperature stimulus for the transition from vegetative to reproductive growth was very important for all of the canola varieties except for the very early maturing variety 1358. Before low temperature induction, the plants will accumulate vegetative biomass. The moderate level of nutrients stored in leaves and un-elongated stems is a benefit for the early stage development of reproductive organs. However, the early planting should be advanced in some limitation. The most important reason is because late maturing canola varieties cannot go through FM differentiation without low temperature induction. However, for the late planting date, a situation similar to that with the very early maturing variety occurred. But the effect of late planting on FM differentiation in canola was more serious than was early planting for the middle and late maturing varieties. For these varieties, low temperature ended during flowering and silique development (**Figures 1** and **3**). Therefore, it is suggested that planting somewhat earlier for middle and late maturing canola varieties is a feasible practice.

In the current study, there were other important considerations as well. First, budding is a key point both for effective flower production and silique formation on all planting dates. FM differentiation is a continuous process once started (**Figure 4**). All the genotype and planting date treatments showed a similar trend in that the number of flowers increased from budding. However, we also found that >20% of the flowers were ineffective based on the ratio of silique formation. Those ineffective flowers or flower primordia represent a considerable waste of resources. Because flowers and flower primordia are not green tissues, they cannot synthesize carbohydrates via photosynthesis. Before the ineffective flower primordia die, they participate in nutrient recycling in different organs such as the leaf and stem (Guiboileu et al., 2010). Their nutrient uptake could possibly affect the effective flower supply during development. Although silique formation is influenced not only by the flower primordium but also by processes such as flower fertilization, the first step is to produce enough flowers. The relatively low ratio of silique formation might be a feedback mechanism for both the degeneration of the FM and the un-fertilized flowers as the temperatures continue to increase. From this perspective, improving the percentage of effective flowers is very meaningful with respect to silique formation, and thus seed yield. In addition to canola, many plants show similar behavior for flower degeneration and drop (Kaska, 1989; Wang et al., 2012). For example, Alburquerque and Egea (2004) showed that low flower bud production and high flower bud drop often resulted in poor yields in apricot. Although the exact regulatory mechanism underlying flower degeneration and drop is not well understood, some hypotheses had been proposed, such as fertilization, nutrient supply, and phytohormone modulation (Buban and Faust, 1982; Pozo, 2001; Sun Y. et al., 2009; Monerri et al., 2011; Boldingh et al., 2016). Recently, Gamer and Lovatt (2016) reported that on average more 70% flowers in avocado trees cannot be fertilized due to the pollen grains unable to germinate and produce pollen tube resulting in ovule degeneration and flower drop. During this process, ABA concentrations in abscising organs were much higher throughout the early to late drop while IAA and isopentenyladenine was much higher at middle drop stage. The result suggested that different phytohormones may function in distinct way on the regulation of flower drop. Another piece of evidence from Boldingh et al. (2016) showed that flowers with successful fruit set had higher carbohydrates and boron content revealing the importance of these matters for flower development to reduce the drop ratio of flowers. In addition to the ABA, jasmonic acid-like

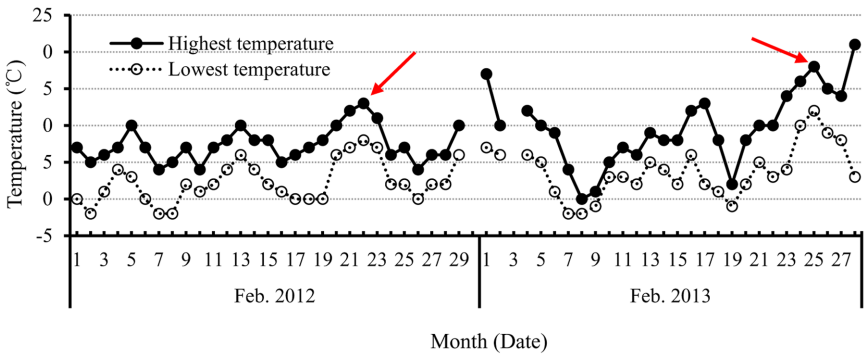


FIGURE 7 | The highest and lowest temperature each day in February in 2012 and 2013, during the period of FM differentiation in canola. The red arrow indicates the unusually high temperatures that occurred during the initiation of FM differentiation to budding stage.

phytohormone was also identified as the similar role on flower and fruit early abscission (Pozo, 2001). Another case, developing flower degeneration occurs because of the competition between different positions of flower. Like other crop such as rice, the flower or seed development is not synchronous. Earlier flower or seed developed usually has stronger opportunity to obtain assimilates compared with younger developing flowers (Fu et al., 2013; Zhang et al., 2014). Furthermore, as temperature increased very quickly after canola flowering (**Figure 1a**), high temperature would together speed up the degeneration of developing flowers (Selak et al., 2014). Therefore, to understand canola flower bud degeneration and flower drop, additional investigations need to be conducted toward improving the effective flower ratio in canola.

The second consideration is how flower numbers, including flower primodia, are associated with the duration of FM differentiation before budding. As discussed above, FM differentiation occurred continuously after budding, but the differentiated flowers were ineffective. In this context, the production of flowers during the stage from FM initiation to budding plays a key role in the total effective numbers of flowers or siliques. Therefore, is there any relationship between flower numbers and the length of the duration from initiation of FM differentiation to budding? Our simple correlation analysis between genotype and flower numbers at the budding stage revealed that only the very early maturing variety showed a significant relationship ($r = 0.88^*$). In addition, correlation analysis of planting date and flower numbers at the budding stage showed there was a correlation with early planting date ($r = 0.92^*$). The non-significant correlation between flower numbers at budding stage and the length of duration from initiation of FM differentiation to budding revealed it should be a complex process because of the complex climatic conditions such as low temperature, tentative high temperature (**Figure 7**), and others. Therefore, optimizing the developmental conditions, including plant resistance to environmental stress and other strategies to prevent or avoid the effects of adverse climatic changes is a vital way to improve canola flower numbers and, hence, siliques.

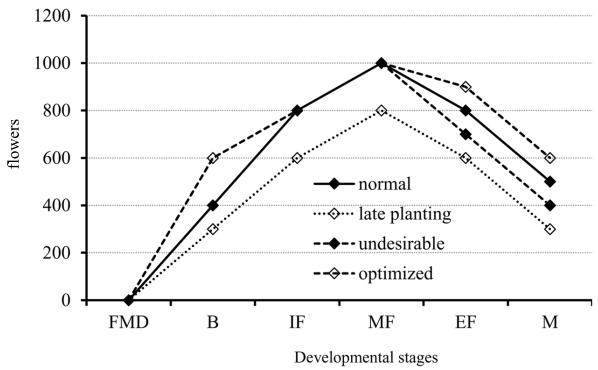


FIGURE 8 | A proposed model for flower development in canola from the initiation of FM differentiation to maturation in response to normal and late planting dates, and other possibly undesirable conditions such as low temperature stress. The flower number was not the real one. The dotted line is for the late planting date that showed fewer flowers at all developmental stages as compared with the normal condition (solid line). The dashed line with open diamonds indicates the potential improvement to maximize flower numbers at the budding, end of flowering, and maturation stages. The dashed line with solid diamonds indicates the potential for flower loss when an undesirable environmental condition occurs. B, budding stage; FMD, floral meristem differentiation stage; IF, initiation of flowering stage; MF, middle flowering stage; EF, end of flowering stage; and M, maturation stage.

Therefore, we derived a model for the initiation of canola floral organs from experiments using varieties that differ with respect to maturity and three planting dates (**Figure 8**). As illustrated in **Figure 8**, there are three important stages: budding, middle flowering, and maturation. The first stage, from FM differentiation to budding, is a fundamental period for flower production. During this time, plants of the middle- and late-maturing varieties often produce fewer flowers (dotted line with open diamonds) and have the potential to maximize flower numbers (dashed line with open diamonds). The second stage is from budding to the middle of the flowering period, and flower numbers for all genotypes have peaked (sometimes the peak will shift to the end of flowering). Most of the flowers produced or the developing flowers are ineffective,

thus, it is presently unknown whether reducing the number of flowers formed during this time will affect final silique formation. The third stage is from the middle of flowering to maturation. During this stage, canola plants suffer from exposure to undesirable conditions, i.e., low temperature stress for the very early variety, and nutrient deprivation and heat stress for all genotypes, which causes them to produce fewer siliques (dashed line with solid diamonds). However, optimizing flower numbers (dashed line with open diamonds) through genetic manipulation, i.e., screening for genotypes with lower ratios of flower degeneration, and agronomic practices such as appropriate application of plant growth regulators, could very well improve yields in canola.

Conclusively, our study showed a clear morphological change of FM developmental process including FM transition from SAM, flower initiation, gynoecium and androecium differentiation, and bud formation in canola. There were no morphological differences between genotypes and planting dates but with significant differences on flower number. Growth days in canola genotypes with different maturity responded to planting date in a delay mode during sowing to enter FM transition while strongly depended on the year and genotype from FM differentiation to bud. Although flower numbers in different genotypes reduced under delayed planting date condition, all the genotypes showed a close number of flowers between budding and harvesting stage. The result suggested an important period from FM differentiation to budding for effective flower production onto fruit set (silique formation). The continuous increasing of flower numbers from budding in five canola genotypes and low ratio of silique numbers to the maximum flowers indicated a great waste of flower investment. Therefore, maximize the flowers before budding and minimize the flower degeneration and drop after budding is a future research and practice direction for canola yield improvement.

AUTHOR CONTRIBUTIONS

YZ performed the floral meristem photographing and wrote the manuscript. DZ design the experiment. HY recorded the growth period at different developmental stages. BL and YF obtained the agronomic trait data. SH design the experiment, data analysis, and manuscript revision.

FUNDING

The work was supported by the earmarked fund for China Agriculture Research System (CARS-13); Special fund by Zhejiang Academy of Agricultural Sciences (2015R16R08E02), Program for Zhejiang Leading Team of S & T Innovation (2011R50026-04 and 2012C12902-1).

ACKNOWLEDGMENT

Great appreciations are given to the editor and reviewers' critical comments on the improvement of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01369

REFERENCES

Adhikari, K. N., Buirchell, B. J., and Sweetingham, M. W. (2012). Length of vernalization period affects flowering in three species. *Plant Breed.* 131, 631–636. doi: 10.1111/j.1439-0523.2012.01996.x

Alburquerque, N., and Egea, L. B. J. (2004). Influence of flower bud density, flower bud drop and fruit set on apricot productivity. *Sci. Hortic.* 102, 397–406. doi: 10.1016/j.scienta.2004.05.003

Bainbridge, K., Sorefan, K., Ward, S., and Leyser, O. (2005). Hormonally controlled expression of the *Arabidopsis* MAX4 shoot branching regulatory gene. *Plant J.* 44, 569–580. doi: 10.1111/j.1365-313X.2005.02548.x

Boldingh, H. L., Alcaraz, M. L., Thorp, T. G., Minchin, P. E. H., Gould, N., and Hormaza, J. I. (2016). Carbohydrate and boron content of styles of 'Hass' avocado (*Persea Americana* Mill.) flowers at anthesis can affect final fruit set. *Sci. Hortic.* 198, 125–131. doi: 10.1016/j.scienta.2015.11.011

Bowman, J. L., and Eshed, Y. (2000). Formation and maintenance of the shoot apical meristem. *Trends Plant Sci.* 5, 110–115. doi: 10.1016/S1360-1385(00)01569-7

Boyhan, G. E., Torrance, R. L., Cook, J., Riner, C., and Hill, C. R. (2009). Sowing date, transplanting date, and variety effect on transplanted short-day onion production. *Horttechnology* 19, 66–71.

Buban, T., and Faust, M. (1982). "Flower bud induction in apple trees: internal control and differentiation," in *Horticultural Reviews*, ed. J. Janick (Hoboken, NJ: John Wiley and Sons, Inc.), 4.

Chen, C., Jackson, G., Neill, K., Wichman, D., Johnson, G., and Johnson, D. (2005). Determining the feasibility of early seeding canola in the north great plains. *Agron. J.* 97, 1252–1262. doi: 10.2134/agronj2005.0004

Djanaguiraman, M., Prasad, P. V. V., Murugan, M., Perumal, R., and Reddy, U. K. (2014). Physiological differences among sorghum (*Sorghum bicolor* L. Moech) under high temperature stress *Environ. Exp. Bot.* 100, 43–54. doi: 10.1016/j.envexpbot.2013.11.013

Ferrándiz, C., Fourquin, C., Prunet, N., Scutt, C. P., Sundberg, E., Trehin, C., et al. (2010). Chapter 1 – Carpel Development. *Adv. Bot. Res.* 55, 1–73. doi: 10.1016/B978-0-12-380868-4.00001-6

Filek, W., and Dubert, F. (1994). Effect of completion of vernalization on generative development of winter rape (*Brassica napus* L. var. oleifera) plants. *J. Agron. Crop Sci.* 172, 29–37. doi: 10.1111/j.1439-037X.1994.tb00156.x

Fu, J., Xu, Y., Yuan, L., Wang, Z., and Yang, J. (2013). Changes in enzyme activities involved in starch synthesis and hormone concentrations in superior and inferior spikelet and their association with grain filling of super rice. *Rice Sci.* 20, 120–128. doi: 10.1016/S1672-6308(13)60116-X

Gamer, L. C., and Lovatt, C. J. (2016). Physiological factors affecting flower and fruit abscission of 'Hass' avocado. *Sci. Hortic.* 199, 32–40. doi: 10.1016/j.scienta.2015.12.009

Gormus, O., and Yucel, C. (2002). Different planting date and potassium fertility effects on cotton yield and fiber properties in the Çukurova region, Turkey. *Field Crops Res.* 78, 141–149. doi: 10.1016/S0378-4290(02)00121-1

Guiboileu, A., Sormani, R., Meyer, C., and Masclaux-Daubresse, C. (2010). Senescence and death of plant organs: nutrient recycling and developmental regulation. *C. R. Biol.* 333, 382–391. doi: 10.1016/j.crvi.2010.01.016

Holman, J., Maxwell, S., Stamm, M., and Martin, K. (2011). Effects of planting date and tillage on winter canola. *Crop Manage.* 10, 1. doi: 10.1094/CM-2011-0324-01-RS

Hormozi, M. A., Asoodar, M. A., and Abdeshahi, A. (2012). Impact of mechanization on technical efficiency: a case study of rice farmer in Iran. *Proc. Econ. Financ.* 1, 176–185. doi: 10.1016/S2212-5671(12)00021-4

Hua, S., Lin, B., Hussain, N., Zhang, Y., Yu, H., Ren, Y., et al. (2014). Delayed planting affects seed yield, biomass production, and carbohydrate allocation in canola (*Brassica napus*). *Int. J. Agric. Biol.* 16, 671–680.

Huang, Y., Zhang, W., Yu, Y., Sun, W., Sun, W., and Chen, J. (2009). *A Primary Assessment of Climate Change Impact on Rice Production in China*. Available at: http://iopscience.iop.org/1755-1315/6/47/472003 [2009]

SAS Institute (2004). *SAS/STAT v.9.1 User's Guide*. Cary, NC: SAS Inst.

Irish, V. F. (2010). The flowering of *Arabidopsis* flower development. *Plant J.* 61, 1014–1028. doi: 10.1111/j.1365-313X.2009.04065.x

Jouannic, S., Lartaud, M., Hervé, J., Collin, M., Orieux, Y., Verdeil, J., et al. (2011). The shoot apical meristem of oil palm (*Elaeis guineensis*;Arecaceae): developmental progress and dynamics. *Ann. Bot.* 108, 1477–1487. doi: 10.1093/aob/mcr019

Kaska, N. (1989). "Bud, flower and fruit drop in citrus and other fruit trees," in *Cell Separation in Plants. Volume H 35 of the serials NATO ASI series, 99*, eds D. J. Osborne and M. B. Jackson (Berlin: Springer-Verlag), 309–321.

Lancashire, P. D., Bleiholder, H., and Van Den Boom, T. (1991). A uniform decimal code for growth stages of crops and weeds. *Ann. Appl. Biol.* 119, 561–601. doi: 10.1111/j.1744-7348.1991.tb04895.x

Li, C., Yue, L., Kou, Z., Zhang, Z., Wang, J., and Cao, C. (2012). Short-term effects of conservation management practices on soil labile organic carbon fractions under a rape-rice rotation in central China. *Soil Till. Res.* 119, 31–37. doi: 10.1016/j.still.2011.12.005

Liang, S., Ye, Q., Li, R., Leng, J., Li, M., Wang, X., et al. (2012). Transcriptional regulations on the low-temperature-induced floral transition in an orchidaceae species, denhrobium nobile: an expressed sequence tags analysis. *Comp. Funct. Genom.* 2012, 757801. doi: 10.1155/2012/757801

Lu, F. (2009). The role agricultural mechanization in the modernization of Asian agriculture: Taiwan's experience. *Eng. Agr. Environ. Food* 2, 124–131. doi: 10.1016/S1881-8366(09)80003-7

Manciocco, A., Calamandrei, G., and Alleva, E. (2014). Global warming and environmental contaminants in aquatic organisms: the need of the etho-toxicology approach. *Chemospere* 100, 1–7. doi: 10.1016/j.chemosphere.2013.12.072

Martínez-Laborda, A., and Vera, A. (2009). "*Arabidopsis* fruit development," in *Annual Plant Reviews : Fruit Development and Seed Dispersal*, Vol. 38, ed. L. Østergaard (Oxford: Wiley-Blackwell).

Monerri, C., Fortunato-Almeida, A., Molina, R. V., Nebauer, S. G., Garcia-Luis, A., and Guardiola, J. L. (2011). Relation of carbohydrate and reserves with the forthcoming crop, flower formation and photosynthetic rate, in the alternate bearing 'Salustiana' sweet orange (*Citrus sinensis* L.). *Sci. Hortic.* 129, 71–78. doi: 10.1016/j.scienta.2011.03.009

Mulyati, Bell, R. W., and Huang, L. (2009). Root pruning and transplanting increase zinc requirements of canola (*Brassica napus* L.). *Plant Soil* 314, 11–24. doi: 10.1007/s11104-008-9701-6

Pickett, F. B., Champagne, M. M., and Meeks-Wagner, D. R. (1996). Temperature-sensitive mutations that arrest *Arabidopsis* shoot development. *Development* 122, 3799–3807.

Pozo, L. V. (2001). Endogenous hormonal status in citrus flowers and fruitlets: relationship with postbloom fruit drop. *Sci. Hortic.* 91, 251–260. doi: 10.1016/S0304-4238(01)00269-2

Selak, G. V., Cuevas, J., Ban, S. G., Pinillos, V., Dumicic, G., and Perica, S. (2014). The effect of temperature on the duration of the effective pollination period in 'Oblica' olive (*Olea europaea*) cultivar. *Ann. Appl. Biol.* 164, 85–94. doi: 10.1111/aab.12082

Sheldon, C. C., Finnegan, E. J., Peacock, W. J., and Dennis, E. S. (2009). Mechanisms of gene repression by vernalization in *Arabidopsis*. *Plant J.* 59, 488–498. doi: 10.1111/j.1365-313X.2009.03970.x

Sindelar, A. J., Roozeboom, K. L., Gordon, W. B., and Heer, W. F. (2010). Corn response to delayed planting in the central great plains. *Agron. J.* 102, 530–536. doi: 10.2134/agronj2009.0311

Sklensky, D. E., and Davies, P. J. (2011). Resource partitioning to male and female flowers of *Spinacia oleracea* L. in relation to whole-plant monocarpic senescence. *J. Exp. Bot.* 62, 4323–4336. doi: 10.1093/jxb/err148

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early flower development in *Arabidopsis*. *Plant Cell* 2, 755–767. doi: 10.1105/tpc.2.8.755

Sorizano, M. A., Orgaz, F., Villalobos, F. J., and Fereres, E. (2004). Efficiency of water use of early plantings of sunflower. *Eur. J. Agron.* 21, 465–476. doi: 10.1016/j.eja.2004.07.001

Souer, E., Van Houwelingen, A., Kloos, D., Mol, J., and Koes, R. (1996). The no apical meristem gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordial boundaries. *Cell* 85, 159–170. doi: 10.1016/S0092-8674(00)81093-4

Stirnberg, P., Zhao, S., Williamson, L., Ward, S., and Leyser, O. (2012). FHY3 promotes shoot branching and stress tolerance in *Arabidopsis* in an AXR1-dependent manner. *Plant J.* 71, 907–920. doi: 10.1111/j.1365-313X.2012.05038.x

Sun, B., Xu, Y., Ng, K., and Ito, T. (2009). A timing mechanism for stem cell maintenance and differentiation in the *Arabidopsis* floral meristem. *Gene. Dev.* 23, 1791–1804. doi: 10.1101/gad.1800409

Sun, Y., Liu, F., Wang, H., and Müller, R. (2009). Effects of ethylene and 1-MCP (1-methylcyclopropene) on bud and flower drop in mini *Phalaenopsis* cultivars. *Plant Grow. Regul.* 59, 83–91. doi: 10.1007/s10725-009-9391-y

Suwa, R., Hakata, H., Hara, H., El-Shemy, H. A., Adu-Gyamfi, J. J., Nguyen, N. T., et al. (2010). High temperature effects on photosynthate partitioning and sugar metabolism during ear expansion in maize (*Zea mays* L.) genotypes. *Plant Physiol. Biochem.* 48, 124–130. doi: 10.1016/j.plaphy.2009.12.010

Takacs, E. M., Li, J., Du, C., Ponnala, L., Janick-Buckner, D., Yu, J., et al. (2012). Ontogeny of the maize shoot apical meristem. *Plant Cell* 24, 3219–3234. doi: 10.1105/tpc.112.099614

Tsimba, R., Edmeades, G. O., Millner, J. P., and Kemp, P. D. (2013). The effect of planting date on maize grain yields and yield components. *Field Crops Res.* 150, 135–144. doi: 10.1016/j.fcr.2013.05.028

Uchida, N., Shimada, M., and Tasaka, M. (2013). ERECTA-family receptor kinases regulate stem cell homeostasis via buffering its cytokinin responsiveness in the shoot apical meristem. *Plant Cell Physiol.* 54, 343–351. doi: 10.1093/pcp/pcs109

Wang, Y., Yang, L., Kobayashi, K., Zhu, J., Chen, C. P., Yang, K., et al. (2012). Investigations on spikelet formation in hybrid rice as affected by elevated tropospheric ozone concentration in China. *Agr. Ecosys. Environ.* 150, 63–71. doi: 10.1016/j.agee.2012.01.016

Williams, L., Gripp, S. P., Xie, M., Christensen, S., and Fletcher, J. C. (2005). Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development* 132, 3657–3668. doi: 10.1242/dev.01942

Wingler, A. (2011). Interactions between flowering and senescence regulation and the influence of low temperature in *Arabidopsis* and crop plants. *Ann. Appl. Biol.* 159, 320–338. doi: 10.1111/j.1744-7348.2011.00497.x

Wollenberg, A. C., and Amasino, R. M. (2012). Natural variation in the temperature range permissive for vernalization in accessions of *Arabidopsis thaliana*. *Plant Cell Environ.* 35, 2181–2191. doi: 10.1111/j.1365-3040.2012.02548.x

Yang, J., and Chen, B. (2014). Global warming impact assessment of a crop residue gasification project-A dynamic LCA perspective. *Appl. Energ.* 122, 269–279. doi: 10.1016/j.apenergy.2014.02.034

Yang, W., Gao, M., Yin, X., Liu, J., Xu, Y., Zeng, L., et al. (2013). Control of rice embryo development, shoot apcial meristem maintance, and grain yield by a novel cytochrome P450. *Mol. Plant* 6, 1945–1960. doi: 10.1093/mp/sst107

Zhang, Z., Zhao, Z., Tang, J., Li, Z., Li, Z., Chen, D., et al. (2014). A proteomic study on molecular mechanism of poor grain-filling of rice (*Oryza sativa* L.) inferior spikelets. *PLoS ONE* 9:e89140. doi: 10.1371/journal.pone.0089140

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhang, Zhang, Yu, Lin, Fu and Hua. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 December 2015
doi: 10.3389/fpls.2015.01164

High-Density SNP Map Construction and QTL Identification for the Apetalous Character in *Brassica napus* L.

*Xiaodong Wang[1,2†], Kunjiang Yu[1†], Hongge Li[1†], Qi Peng[1], Feng Chen[1,3], Wei Zhang[1], Song Chen[1], Maolong Hu[1] and Jiefu Zhang[1,2]**

[1] Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture, Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [2] Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing, China, [3] Provincial Key Laboratory of Agrobiology, Jiangsu Academy of Agricultural Sciences, Nanjing, China

OPEN ACCESS

Edited by:
Juan Francisco Jimenez Bremont, Instituto Potosino de Investigacion Cientifica y Tecnologica, Mexico

Reviewed by:
Estefanía Carrillo, Instituto Nacional de Investigaciones Agropecuarias, Ecuador
Javier Sanchez, University of Zürich, Switzerland

***Correspondence:**
Jiefu Zhang
jiefu_z@163.com

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *10 September 2015*
Accepted: *07 December 2015*
Published: *23 December 2015*

Citation:
Wang X, Yu K, Li H, Peng Q, Chen F, Zhang W, Chen S, Hu M and Zhang J (2015) High-Density SNP Map Construction and QTL Identification for the Apetalous Character in Brassica napus L.. Front. Plant Sci. 6:1164. doi: 10.3389/fpls.2015.01164

The apetalous genotype is a morphological ideotype for increasing seed yield and should be of considerable agricultural use; however, only a few studies have focused on the genetic control of this trait in *Brassica napus.* In the present study, a recombinant inbred line, the AH population, containing 189 individuals was derived from a cross between an apetalous line 'APL01' and a normally petalled variety 'Holly'. The *Brassica* 60 K Infinium BeadChip Array harboring 52,157 single nucleotide polymorphism (SNP) markers was used to genotype the AH individuals. A high-density genetic linkage map was constructed based on 2,755 bins involving 11,458 SNPs and 57 simple sequence repeats, and was used to identify loci associated with petalous degree (PDgr). The linkage map covered 2,027.53 cM, with an average marker interval of 0.72 cM. The AH map had good collinearity with the *B. napus* reference genome, indicating its high quality and accuracy. After phenotypic analyses across five different experiments, a total of 19 identified quantitative trait loci (QTLs) distributed across chromosomes A3, A5, A6, A9 and C8 were obtained, and these QTLs were further integrated into nine consensus QTLs by a meta-analysis. Interestingly, the major QTL *qPD.C8-2* was consistently detected in all five experiments, and *qPD.A9-2* and *qPD.C8-3* were stably expressed in four experiments. Comparative mapping between the AH map and the *B. napus* reference genome suggested that there were 328 genes underlying the confidence intervals of the three steady QTLs. Based on the Gene Ontology assignments of 52 genes to the regulation of floral development in published studies, 146 genes were considered as potential candidate genes for PDgr. The current study carried out a QTL analysis for PDgr using a high-density SNP map in *B. napus*, providing novel targets for improving seed yield. These results advanced our understanding of the genetic control of PDgr regulation in *B. napus*.

Keywords: *Brassica napus* L., apetalous, single nucleotide polymorphism, high-density map, quantitative trait locus, recombinant inbred line

INTRODUCTION

Oilseed rape (*Brassica napus* L., AACC, 2n = 38) is a widely planted oil crop worldwide. Rapeseed oil is not only a desirable edible oil, but is also used as a biofuel in many parts of the world (Hoekman, 2007). Additionally, oil-extracted meal from *Brassica* seeds is an excellent protein source for animal feed (Linnemann and Dijkstra, 2002). As the global demand for rapeseed products is continuously increasing, developing a high-yield variety is a main goal of *B. napus* breeding programs. An effective approach is to seek the morphological ideotype (Virk et al., 2004), and apetalous genotypes are of particular interest in breeding programs (Habekotté, 1997; Jiang and Becker, 2003).

The apetalous trait was first reported in a naturally occurring mutant of turnip (*B. campestris* L.) in *Brassica* (Ramanujam, 1940), and Buzza (1983) first detected an apetalous mutant in a spring oilseed rape. Since then, other apetalous flowers in *Brassica* species have been discovered or bred. There are several advantages to apetalous rape in yield. First and foremost, photosynthesis in cultivars without petals is more efficient, with the thick and brightly colored flowers preventing *Brassica* oilseeds from efficiently using solar energy (Jiang, 2007). The petals at the top layer of the normal flower type were reported to reflect or absorb up to 60% of incoming radiation (Mendham et al., 1981). Additionally, the apetalous cultivars have higher yield potentials than the normal type. The petal is not a photosynthesizing organ, but it consumes considerable amounts of photosynthesized assimilates during its formation and respiration (Jiang, 2007). Mendham et al. (1991) revealed that the yield of apetalous lines was higher than normal petaled cultivars. Finally, the apetalous type of rapeseeds might have a lower rate of infection from diseases distributed by petals, such as *Sclerotinia sclerotiorum*. Deciduous petals can transmit the *Sclerotinia* pathogen to healthy tissue, whereas the ascospores that land directly on the leaf surface do not germinate (Jamaux and Spire, 1999). Compared with normal petaled controls, apetalous genotypes have a much lower incidence and severity of *Sclerotinia* infection (Lefol and Morrall, 1996; Zhao and Wang, 2004). Moreover, a multitude of other diseases, such as *Botrytis cinierea* and *Peronospora parasitica*, may be distributed by petals (Lefol and Morrall, 1996). In summary, genotypes with apetalous flowers are a component of the high-yielding ideotype.

In *B. napus*, various genetic models of the apetalous trait, with different origins, are documented in many literatures. One study found that petalous flower development was controlled by one gene locus that exhibited incomplete dominance over apetalous flower development (Zhao and Wang, 2004). Most of the other studies revealed that the apetalous character in *B. napus* was regulated by recessive genes, possibly by two to four loci (Buzza, 1983; Lu and Fu, 1990; Kelly et al., 1995; Chen et al., 2006a; Zhang et al., 2007b,c). Generally, these loci independently control flower morphology; however, epistatic interactions between recessive alleles were also identified (Fray et al., 1997). In addition, the apetalous character is governed by the interaction of cytoplasmic and nuclear genes (Jiang and Becker, 2003). Although much attention has been paid to the inheritance of the apetalous character in *B. napus*, questions concerning the genetic basis remain open. Plants having less than a 10% petalous degree (PDgr) were considered apetalous (Buzza, 1983), but the distribution of PDgr in segregation generation is consecutive and should be treated as a quantitative trait (Zhang et al., 2007a). Quantitative trait loci (QTLs) mapping is a preliminary step and an effective approach to unravel the genetic architecture of complex quantitative traits and to identify QTLs for knowledge-based breeding (Mauricio, 2001). Fray et al. (1997) reported that five restriction fragment length polymorphism markers were significantly associated with one of two *stamenoid petal* loci, one each on A4 and C4. One random amplified polymorphic DNA marker that was tightly linked to a petal-controlled gene in *B. napus* was identified by Tan et al. (2003). Using a bulked segregant analysis approach, one sequence-related amplified polymorphism and one amplified fragment length polymorphism marker mapped on A4 were found to be linked to the gene controlling the petal-loss trait (Chen et al., 2006b). Based on a genetic map containing 219 markers, Zhang et al. (2007a) identified four QTLs, which were located on chromosomes A5, A6, A8, and C5, associated with the apetalous phenotype.

High-density maps could increase the precision of QTL localization and the estimation of QTL effects in biparental populations (Stange et al., 2013). Since the first molecular linkage map in *B. napus* was reported by Landry et al. (1991), various types of populations have been constructed for mapping QTLs associated with seed oil content, seed fatty acid concentrations, flowering time, seed yield and yield-related traits (Qiu et al., 2006; Quijada et al., 2006; Udall et al., 2006; Long et al., 2007; Basunanda et al., 2010; Chen et al., 2010; Cai et al., 2012; Ding et al., 2012; Wang et al., 2013; Wang et al., 2015a). However, most of these genetic linkage maps were constructed based on PCR markers with low densities. Single nucleotide polymorphisms (SNPs) are the most frequent polymorphism in the genomes of crops (Vignal et al., 2002), and they have been widely used in rice (Huang et al., 2010), wheat (Mochida et al., 2003), and maize (Ching et al., 2002). In *B. napus*, SNPs were also used for high-density genetic map construction and the fine mapping of important genes. Delourme et al. (2013) developed an integrated genetic map, which was comprised of 5,764 SNPs and 1,603 PCR markers, with a genetic length of 2,250 cM. Based on the 6 K SNP array harboring 5,306 probes for *B. napus*, Raman et al. (2014) constructed a genetic linkage map covered 2,514.8 cM, including 613 SNPs and 228 non-SNPs, and Cai et al. (2014) constructed a genetic map containing 2,115 markers (1,667 SNPs and 448 SSRs), with a length of 2,477.4 cM. Chen et al. (2013) constructed a SNP bin map containing 8,780 SNP loci and a presence/absence variation map containing 12,423 dominant loci. In 2012, the *Brassica* 60 K SNP BeadChip Array comprised of 52,157 SNP loci was produced (Snowdon and Iniguez Luy, 2012; Edwards et al., 2013), which was developed by an international consortium using preferentially single-locus SNPs contributed from genomic and transcriptomic sequencing in genetically diverse *Brassica* germplasm (Liu et al., 2013). This paved the way for the high-throughput and cost-effective construction of a high-density genetic *B. napus* map. Using the 60 K SNP BeadChip Array,

Liu et al. (2013) constructed a linkage map containing 9,164 SNP markers covering 1,832.9 cM, and mapped the major QTL for seed color corresponding to a physical region of 620 kbp. Zhang et al. (2014) constructed a map covering a length of 2,139.5 cM with average distance of 1.6 cM between adjacent markers. Nevertheless, to the best of our knowledge, a QTL analysis for the apetalous trait using high-density SNP genetic linkage map in *B. napus* has not been performed.

The objectives of the present study were to: (1) construct a high-density genetic map using the *Brassica* 60 K Infinium SNP array and SSRs; and (2) investigate the QTLs for PDgr in *B. napus* across five experiments. The results will provide information useful for understanding the genetic control of apetalous in *B. napus*, and the major QTLs will lay a foundation for use in breeding programs to develop a variety with agronomic traits of interest for rapeseed production.

MATERIALS AND METHODS

Plant Materials

The *B. napus* segregating recombinant inbred line (RIL) population used in this study was derived from a cross between 'APL01' and 'Holly' using the single seed descent method. The parent 'APL01' is an apetalous line, developed at the Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing, China. 'APL01' was selected from F_6 generation of crosses between apetalous (Apeatlous No. 1) and normal petalous (Zhongshuang No. 4) rapeseed in 1998 (Zhang et al., 2002). 'Apeatlous No. 1' was bred from F_8 generation of crosses between China rapeseed cultivar with smaller petals (SP103) and *B. rapa* variety with lower petals (LP153). 'Zhongshuang No. 4' was developed at Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China. Except for PDgr, other agronomic traits in 'APL01' are quite normal. At early flowering stage, 'APL01' is absolutely apetalous, however, there may be one cripple petal only in a few flowers at late flowering stage. The genotype 'Holly' is a normally and completely petaled variety. The two parents showed the similar flowering time, which recorded from the sowing day to the day when the first flower had opened on half of the plants in the plot. A total of 550 F_9 RIL lines were developed in 2014, and then a subset of 189 lines was randomly selected to compose the mapping population for the genetic linkage map construction. These were named the AH RIL population.

Field Trials and PDgr Measurements

The AH population, together with the two parents, was tested in five experiments. The materials were planted in a winter rapeseed area, Dali of Shaanxi Province (coded DL), in northwest China for one year (September—May of 2014–2015), and a semi-winter rapeseed area, Nanjing of Jiangsu Province (coded NJ), in eastern China for four years (September—May of 2011–2012, 2012–2013, 2013–2014, and 2014–2015). Year-location combinations were treated as experiments, for example, 14NJ indicates the experiment was conducted during 2014–2015 at the Nanjing location. The field experiments were conducted in a randomized complete block design with two replications in both NJ and DL. The experimental unit was a two-row plot with 20 plants per row and 40 cm between the rows. The field management followed the common agricultural practices.

At least five representational plants of each plot were selected to measure PDgr during the flowering stage. The percentage of petals of an individual plant was determined by counting the number of petals on the first 25 flowers to open (Buzza, 1983). The PDgr was calculated by the following formula as described by Buzza (1983):

$$\text{PDgr (\%)} = \left(\sum_{i=1}^{n} Pi/4n\right) \times 100\%,$$

in which P represents the number of petals for each flower that was counted, with a range of numerical value of 0–4, and n is the total number of flowers we investigated, $n \geq 50$.The average of the PDgr in each RIL line was used as raw data in the analysis.

Statistical Analysis

Basic statistical analyses of PDgr were performed using SPSS 18.0 software (SPSS Inc., Chicago, IL, USA). The software package SEA-G3DH, with the mixed major gene and polygene inheritance model, was used to analyze the inheritance of the PDgr character in the AH population (Cao et al., 2013). The best fitting model was selected from 39 different models that were included in the software package according to Wang et al. (2015b). Because the PDgr was expressed as a percentage and did not fit the normal distribution model, all data were subjected to an arcsine transformation for the genetic model analysis. Genetic parameters were estimated using the best fitting model with the default settings in the software.

SNP and SSR Marker Analysis

The genotypes of the AH RIL population and two parental lines were analyzed using the *Brassica* 60 K SNP BeadChip Array, which successfully assays 52,157 Infinium Type II SNP loci in *B. napus*. This array was developed by the international *Brassica* SNP consortium in cooperation with Illumina Inc. San Diego, CA, USA. DNA sample preparation, hybridization to the BeadChip, washing, primer extension and staining were strictly carried out according to the Infinium HD Assay Ultra manual. Imaging of the arrays was performed using an Illumina HiSCAN scanner. Allele calling for each locus was performed using the GenomeStudio genotyping software v2011 (Illumina, Inc.). SNP markers used the names that were assigned by GenomeStudio, such as "Bn-A01-p25032772". SSR primer pairs prefixed "CB" and "BRAS" were published by Piquemal et al. (2005); "Na", "Ol" and "Ra" were developed by Lowe et al. (2004); "MR" were published by Uzunova and Ecke (1999) and "BnGMS" were developed by Cheng et al. (2009).

Construction of the Genetic Linkage Map and Alignment of the *B. napus* Reference Genome

All SNPs that were polymorphic between 'APL01' and 'Holly', as well as having less than 5% missing data, were used for the

genetic linkage map construction. Of the 52,157 SNPs in the array, 17,414 SNPs met the above requirements and were selected for further analysis. SNP marker pairs with no recombination were classified into one genetic bin (one bin corresponded to all of the markers having the same genotype scoring data) using a Perl script. Then, the selected 17,414 SNPs were grouped into 3,422 SNP-bins, containing 1 to 2,079 SNPs in each bin. Combined with 81 polymorphic SSR markers between the two parents, 3,503 loci (3,422 SNP-bins and 81 SSRs) were subsequently applied in map construction using JoinMap software Version 4.0 (Van Ooijen, 2006). Centimorgan (cM) distances were calculated by the Kosambi function for map distance (Kosambi, 1943). Markers with a mean chi-square value $\geq$ 3.0 were excluded in all genetic groups to ensure the high quality of the map (Wang et al., 2013).

In addition, the probe sequences of the SNPs that assigned the A and C sub-genomes of *B. napus*, were queried using the BLAST algorithm against the *B. napus* reference genome sequence to locate chromosomal positions with highly stringent parameters (*E* value $\leq$ 1e-10) (Chalhoub et al., 2014). Alignments between the SNP bin map and the *B. napus* reference genome were used to validate the quality of the genetic map. If a locus was mapped to multiple paralogous positions in the *B. napus* reference genome, only the location that corresponded to the particular linkage group of the locus was selected for the collinearity analysis.

QTL Detection and Meta-Analysis

The software Windows QTL Cartographer 2.5 with a composite interval mapping model was used to estimate putative QTLs with additive effects (Zeng, 1994; Wang et al., 2007). The walking speed was set to 2 cM, and a window size of 10 cM with five background cofactors was used. The LOD threshold (2.8–3.1) for detection of significant QTLs was set by a 1,000-permutation test based upon a 5% experiment-wise error rate, and these QTLs were termed 'identified QTLs'. Identified QTLs that were detected in different experiments with overlapping confidence intervals (CIs), may have been one single QTL. Then, identified QTLs were integrated into consensus QTLs using a meta-analysis method with the BioMercator V4.2 program (Arcade et al., 2004). If an identified QTL had no overlapping CIs with others, then it was also regarded as a consensus QTL (Wang et al., 2013).

The identified and consensus QTL nomenclature was based on the descriptions of McCouch et al. (1997) with minor modifications. For identified QTLs, a designation begins with the abbreviation "*iq*" (identified QTL), follow by the experiment and linkage group (A1–A10, C1–C9). If there was more than one identified QTL obtained in a linkage group, a serial number was added. For example, *iq12NJ.A9-2* indicates the second identified QTL for PDgr on A9 in the 12NJ experiment. For consensus QTLs, a designation begins with the abbreviation "*qPD*" (*q*, QTL; *PD*, petalous degree). For example, QTL *qPD.A9-2* indicates the second consensus QTL for PDgr on the A9 linkage group. In addition, candidate genes were identified by comparative mapping between the AH map and the *B. napus* reference genome based on the probe sequence of the SNPs. If the physical positions of aligned genes fell into the CI of a consensus QTL, then the orthologous candidate genes were assumed to be associated with the target QTL (Ding et al., 2012).

Results

Phenotypic Variation and Genetic Analysis for PDgr

'APL01' and 'Holly' are the two parental lines whose repeatability is very good in the five experiments. The mean values $\pm$ SE were 0.02% $\pm$ 0.03 and 99.99% $\pm$ 0.01 for 'APL01' and 'Holly', respectively. There was a wide range of variations and consecutive distributions in the AH population (**Figure 1**), suggesting that PDgr was governed by multiple genes. However, the phenotypic values did not fit the normal distribution in the majority of the RIL lines, mostly petalled, indicating that PDgr might be determined mainly by the major genes in *B. napus*.

The mixed major gene and polygene inheritance model has been used to analyze the two parents and AH population for PDgr in the 11NJ and 12NJ experiments (Li et al., 2014). The PDgr of the AH population was controlled by the two additive major genes and the additive polygene model (MX2-Additive-A model) (Li et al., 2014). Using the same method, the best fitness genetic models for PDgr in 13NJ, 14NJ, and 14DL were analyzed in the present study. The MX2-Additive-A model was also the best fitness genetic model in the 13NJ and 14NJ experiments. However, the genetic model MX2-EA-A, which mixed two equal additive major genes with the additive polygene model, was the most suitable model for the 14DL experiment. The heritabilities of major genes ranged from 68.52% to 88.01% in the

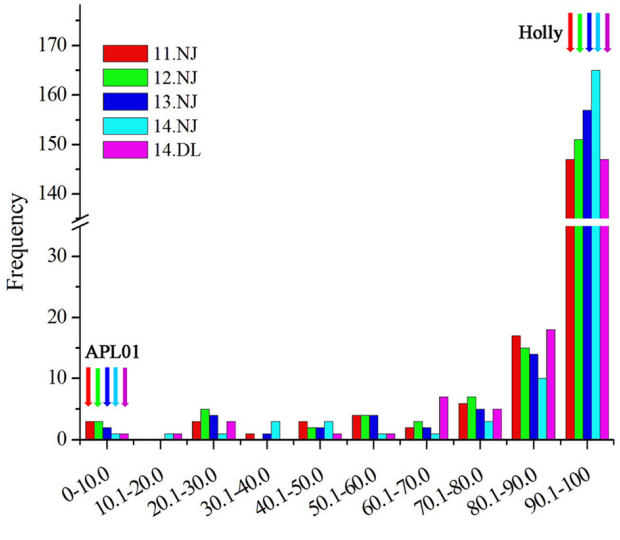


FIGURE 1 | Distribution of the petalous degree (PDgr) of the AH RIL population in five experiments. The units of the *x*-axis are the percentage of the PDgr, and the units of the *y*-axis are the number of lines. PDgr in different experiments was discriminated using different colored boxes (11NJ, red; 12NJ, green; 13NJ blue; 14NJ, cyan; 14DL, magenta). The color arrows represent the location of the parental lines.

five experiments (**Table 1**), significantly higher than that of the polygenes (11.99–31.48%), indicating that PDgr in *B. napus* was determined by the combination of major genes and polygenes, but mainly by the major genes.

High-Density SNP Map Construction

Among the 3,503 loci (3,422 SNP-bins and 81 SSRs) that were used for map construction, 2,755 SNP-bins and 57 SSRs were assigned to 19 linkage groups, including 1,686 loci in the A sub-genome (A1–A10) and 1,126 in C sub-genome (C1–C9; **Table 2**, Supplementary Table S1). The number of SNP markers varied considerably across the different bins, ranging from 1 to 317, and 11,458 SNPs involved in the 2,812 loci were assigned to the genetic map (**Table 2**; **Figure 2**). A8 has the lowest number of SNPs with only 139 SNPs spanning 76.77 cM, and C2 has the most SNPs with 1,788 SNPs spanning 112.08 cM (**Figure 2**).

The high-density map had a total length of 2,027.53 cM with an average marker interval of 0.72 cM, covering 1,033.31 and 994.22 cM of the A and C sub-genomes, respectively (**Table 2**). The average linkage group lengths of the A and C sub-genomes were similar at 103.3 and 110.5 cM, respectively. However, the lengths of each group showed great differences, ranging from 76.09 (A4) to 134.91 cM (A1) in the A sub-genome and 65.10 (C1) to 157.02 cM (C4) in the C sub-genome. In addition, no chromosome in the genetic map displayed gaps of more than 20 cM, while C5 showed the largest gap of 16.70 cM between Bn-scaff_16082_1-p33791 and Bn-scaff_15712_10-p52253 (Supplementary Table S1).

Alignment of the SNP Linkage Map to the *B. napus* Reference Genome

The probe sequences of all 2,755 SNP-bins that mapped to the 19 linkage groups were aligned to the *B. napus* reference genome to validate the genetic linkage maps (**Figure 3**). The results showed that 2,350 loci produced successful BLAST hits in the *B. napus* database, accounting for 85.30% of the 2,755 SNP-bins (85.63% for the A sub-genome and 84.80% for the C sub-genome).

Alignments indicated that the linkage map constructed in the present study had good collinearity with the *B. napus* reference genome sequence (**Figure 3**), suggesting the high quality of the AH RIL map. However, several inconsistencies on A1, A3, A7, and A8 were detected between the map and the *B. napus* reference genome. For example, a large inconsistency involving 80 SNP-bins and spanning a region from 100.82 to 133.75 cM (16.27 Mb of physical region) on A1 might be caused by the existence of paralogous sequences. The inconsistency from the 34.36 to 45.37 cM on the A3 chromosome, which includes 38 SNP-bins and corresponds to ∼1.23 Mb of the physical interval, may have been caused by the presence of partial homologous sequences or fragment duplications. In addition, an inversion including 41 loci and spanning from 64.33 to 71.49 cM (2.49 Mb) was identified on the A7 chromosome, and an inversion with 20 loci from 4.09 to 12.40 cM (2.54 Mb) was also identified on A8.

The total genome size of *B. napus* is estimated to be 1,130 Mb, and ∼645.4 Mb of the genome assembly was collectively comprised by the scaffolds (Chalhoub et al., 2014). The A sub-genome of the AH map showed good coverage of the reference *B. napus* genome, representing 93.72% of the genome assembly length, while the C sub-genome showed a much lower coverage of 76.91% (**Table 2**). The main reason was that three linkage groups, C1, C3, and C5, only represent 33.45, 43.70, and 34.17%, respectively, of the corresponding chromosomes (**Table 2**; **Figure 3**). The remaining six linkage groups of the C sub-genome account for 92.21to 99.71% of the genome assembly length.

QTL Detection and Meta-Analysis for PDgr

Phenotypic data of PDgr in the AH population were obtained from the five different environments. QTLs for PDgr were analyzed based on the high-density SNP map, and then identified QTLs were integrated into consensus QTLs. Detailed information on identified and consensus QTLs are summarized in **Table 3**.

A total of 19 identified QTLs distributed across A3 (1 QTL), A5 (1 QTL), A6 (1 QTL), A9 (6 QTLs), and C8 (10 QTLs) chromosomes were detected in the present study (**Table 3**). These QTLs have additive effects ranging from -6.09 to 6.63, and singly explaining 5.08%–11.29% of the estimated phenotypic variation (PV). The number of identified QTLs detected in different experiments was quite different, ranging from two (14NJ) to five (11NJ). Among them, up to 10 identified QTLs were distributed on C8, implying that the major genes for PDgr might exist on the C8 chromosome. Eighteen of the 19 identified QTLs had negative additive effects, suggesting that the normally petalled parent 'Holly' contributed favorable alleles for increasing PDgr. Only one identified QTL, *iq12NJ.A3*, had the positive additive effect of 6.63, indicating that the positive alleles for higher phenotypic values were inherited from the apetalous parent 'APL01'.

There were 13 identified QTLs with overlapping CIs, and these were further integrated into three consensus QTLs using a meta-analysis method (**Table 3**; **Figure 4**). As a result, the average CIs of these QTLs were reduced from 5.55 to 1.94 cM, which significantly increased the accuracy of the estimated positions of the meta-QTL. The other six non-overlapping QTLs were also considered as consensus QTLs. In total, nine consensus QTLs for PDgr were obtained in the present study. Among these consensus QTLs, one QTL (*qPD.C8-2*) with the closely linked marker Bn-scaff_18275_1-p1278049 of 0.15 cM, was consistently detected in all of the five experiments and had additive values in the range of −0.69 to −3.77 (**Table 3**). According to the description of Shi et al. (2009), if a consensus QTL presents at least once with PV ≥ 20% or at least twice with PV ≥ 10%, then the QTL can be regarded as a major QTL. The QTL *qPD.C8-2* with PV ≥ 10% in 11NJ, 12NJ and 13NJ (10.47, 11.29, and 10.24%, respectively) was the major QTL, which might harbor major genes responsible for PDgr. In addition, two QTLs, *qPD.A9-2* and *qPD.C8-3*, were expressed steadily in four experiments, with linked markers of Bn-A09-p29172005 and Bn-scaff_17227_1-p700248 and explained 6.05%–7.06% and 6.12%–8.40% of PV, respectively. The genes associated with these QTLs controlling PDgr may be less affected by environment,

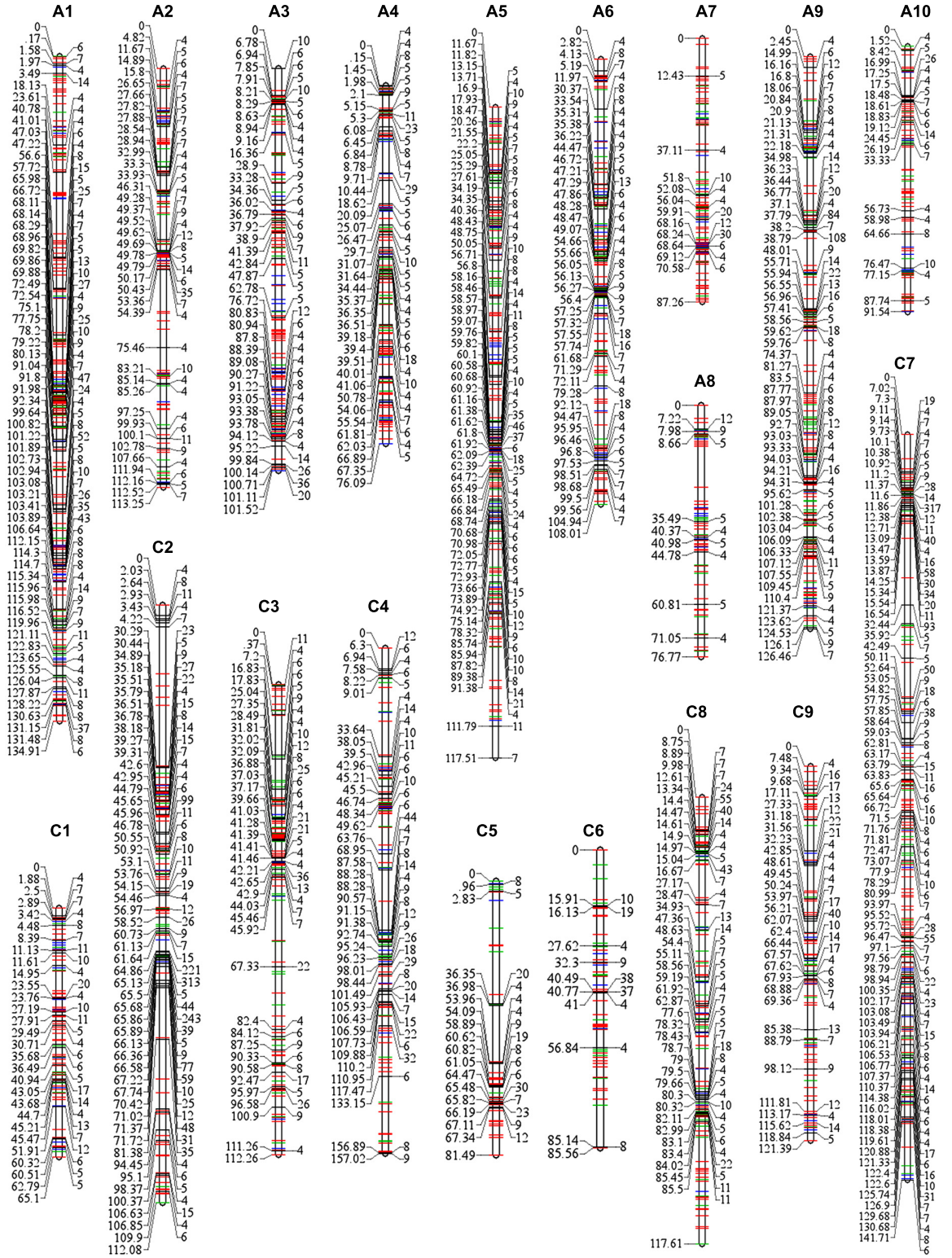


FIGURE 2 | SNPs distribution in each linkage group of the map constructed using 189 RIL individuals. The 19 linkage groups are represented by vertical bars, designated as A1–A10 in the A sub-genome and C1–C9 in the C sub-genome. The number of SNPs in each bin is listed on the right side of the linkage groups, while the positions of the bins are shown on the left side of the linkage groups. Bins with less than three SNPs were represented by colored lines (red lines, bins involve one SNP; green lines, two SNPs; and blue lines, three SNPs), for simplicity the numbers and positions are not shown. Full details are provided in Supplementary Table S1.

TABLE 1 | Genetic parameters estimated in the MX2-Additive-A or MX2-EA-A model in the AH population.

First order	Experiments[b]					Second order	Experiments				
Parameter[a]	11NJ	12NJ	13NJ	14NJ	14DL	parameter[c]	11NJ	12NJ	13NJ	14NJ	14DL
d(da)	−8.24	−9.31	−9.25	−10.85	−18.41	σ^2_p	332.53	235.52	333.76	286.01	185.5
Db					−18.41	σ^2_{mg}	287.11	198.41	290.57	251.65	127.1
I	−21.17	−20.9	−20.44	−20.35		σ^2_{pg}	44.87	36.56	43.18	34.29	58.38
[*d*]	−28.14	−26.03	−26.48	−23.3	0	σ^2_e	0.55	0.56	0.01	0.07	0.02
						h^2_{mg}(%)	86.41	84.34	87.06	88.01	68.52
						h^2_{pg}(%)	13.51	15.54	12.94	11.99	31.48

DL, Dali; NJ, Nanjing; 11, 12, 13 and 14 indicate the years 2011, 2012, 2013, and 2014, respectively. [a]*d: The additive effect of a major gene; i: The epistasis effect of a major gene; [d]: The additive effect of polygene.* [b]*The best fitness genetic model in 11NJ, 12NJ, 13NJ and 14NJ experiments is the MX2-Additive-A model, while the MX2-EA-A model was the best fitness genetic model in the 14DL experiment. Genetic parameters estimated in 11NJ and 12NJ were reported by Li et al. (2014).* [c]σ^2_p*, Population variance;* σ^2_{mg}*, Major gene variance;* σ^2_{pg}*, Polygene variance;* σ^2_e*, Environmental variance;* h^2_{mg} *(%), Heritability of the major genes;* h^2_{pg} *(%), Heritability of the polygene.*

which is consistent with the results of the genetic analysis, which indicated that PDgr in the AH population has a high heritability (**Table 1**). A striking finding was that, except for the three abovementioned QTLs, no other QTL was detected in multiple experiments, and the remaining six QTLs were environment-specific QTLs detected only in one environment (**Table 3**).

Identification of Candidate Genes Responsible for PDgr in the AH Population

As mentioned above, *qPD.A9-2*, *qPD.C8-2,* and *qPD.C8-3* were the three steady QTLs that showed significant effects in four, five, and four environments, respectively (**Table 3**). The three QTLs span regions of 75.97–76.75 cM, 15.69–19.7 cM, and 26.69–27.72 cM, with physical regions of 0.35, 2.65, and 0.26 Mb on chromosomes A9, C8, and C8, respectively. Based on the comparative mapping between the AH map and the *B. napus* reference genome, 58, 220, and 50 genes, respectively, were identified underlying the CIs of the three QTLs (Supplementary Table S2A). The other six consensus QTLs were only detected in certain geographical regions, suggesting that they were easily influenced by the environment, and genes underlying these QTLs were not analyzed in the present study.

Fifty-two genes regulating floral development were collected from 55 published studies that have been performed mainly in *Arabidopsis* (Supplementary Table S2B). These genes were further annotated with gene ontology (GO) terms, which were classified into three categories: biological process (BP, **Figure 5A**), cellular component (CC, **Figure 5B**), and molecular function (MF, **Figure 5C**). The 52 genes were classified into 92 functional groups, with the number of genes in each GO term ranging from 1 to 50 (Supplementary Table S2C). In the BP category, the most abundant GO terms were GO:0006355 (regulation of transcription, DNA-templated) and GO:0006351 (transcription, DNA-templated). In the CC category, GO:0005634 (nucleus) was the most abundant, followed by GO:0005737 (cytoplasm). In the MF category, GO:0005515 (protein binding) and GO:0003700 (sequence-specific DNA binding transcription factor activity) were the major GO terms.

Gene ontology assignments were also used to classify the functions of the 328 genes that underlie the three steady QTLs. The results showed that 146 genes were distributed under 38 of the 92 GO terms (**Figure 5**, Supplementary Table S2C), and thus are potential candidate genes for PDgr in *B. napus*. Among these, 34 (14 and 20, respectively), 25 (11 and 14, respectively), and 64 (57 and 7, respectively) genes were classified in the top two GO terms in the BP, CC and MF categories to which the published genes for petal development were assigned, respectively (**Figure 5**, Supplementary Table S2C). Some candidate genes were assigned to more than one category. For example, *BnaC08g10960D* and *BnaC08g14030D*, which underlie the CIs of *qPD.C8-2* and *qPD.C8-3*, respectively, were assigned to all three categories. In addition, seven, eight and five candidate genes were simultaneously assigned to BP and CC, BP and MF, CC and MF categories, respectively.

DISCUSSION

The apetalous genotype is a morphological ideotype for increasing seed yield compared with the fully petalled variety in *B. napus*, which improves light penetration through the floral canopy and lowers the incidence of *S. sclerotiorum* infection (Mendham et al., 1981, 1991; Jamaux and Spire, 1999; Jiang, 2007). Having suitable parental genotypes is critical for identifying the inheritance and the genetic bases of a character; however, the lack of completely apetalous genetic resources has been a limiting factor for understanding its genetic control in *B. napus*. Several apetalous flowers in *B. napus* were used to investigate the genetic models for the apetalous character (Buzza, 1983; Lu and Fu, 1990; Kelly et al., 1995; Zhao and Wang, 2004; Chen et al., 2006a; Zhang et al., 2007b,c). In the present study, an apetalous line 'APL01' and a normally petalled variety 'Holly' were used to construct the AH RIL population and perform a genetic analysis of the apetalous trait. The results showed that PDgr in *B. napus* was controlled by the combination of two major genes and polygenes, which was consistent with previous studies (Buzza, 1983; Kelly et al., 1995; Jiang and Becker, 2003; Chen et al., 2006a; Zhang et al., 2007b). Other genetic models

TABLE 2 | Summary of the high-density SNP map based on the AH RIL population.

Chromosome	SNP-bins	SNPs	SSRs	SNPs+SSRs	Bins+SSRs	Length (cM)	Density	No. of loci assigned to *B. napus* genome	Genome coverage (Mb)	Total length of *B. napus* genome (Mb)	% of genome coverage
A1	270	1008	7	1015	277	134.91	0.49	210	0.20–23.24 (23.04)	23.27	99.01
A2	151	430	0	430	151	113.25	0.75	126	0.06–24.79 (24.73)	24.79	99.76
A3	184	528	2	530	186	101.52	0.55	171	2.12–28.99 (26.87)	29.77	90.26
A4	141	413	1	414	142	76.09	0.54	119	0.74–19.08 (18.34)	19.15	95.77
A5	226	836	5	841	231	117.51	0.51	202	0.09–21.62 (21.53)	23.07	93.32
A6	175	493	4	497	179	108.01	0.60	148	1.21–23.84 (22.63)	24.40	92.75
A7	118	260	3	263	121	87.26	0.72	100	1.79–22.13 (20.34)	24.01	84.71
A8	68	139	4	143	72	76.77	1.07	65	0.73–18.86 (18.13)	18.96	95.62
A9	212	832	2	834	214	126.46	0.59	172	4.14–33.83 (29.69)	33.87	87.66
A10	111	269	2	271	113	91.54	0.81	105	0.26–17.37 (17.11)	17.40	98.33
C1	100	326	2	328	102	65.10	0.64	87	0.08–13.07 (12.99)	38.83	33.45
C2	146	1788	0	1788	146	112.08	0.77	116	0.52–45.79 (45.27)	46.22	97.94
C3	136	499	2	501	138	112.26	0.81	127	0.002–26.47 (26.47)	60.57	43.70
C4	124	585	5	590	129	157.02	1.22	104	0.002–48.79 (48.79)	48.93	99.71
C5	46	215	2	217	48	81.49	1.70	40	0.15–14.91 (14.76)	43.19	34.17
C6	63	212	1	213	64	85.56	1.34	50	2.86–37.19 (34.33)	37.23	92.21
C7	229	1612	8	1620	237	141.71	0.60	199	0.63–44.32 (43.69)	44.77	97.59
C8	153	587	2	589	155	117.61	0.76	136	0.80–37.98 (37.18)	38.48	96.62
C9	102	426	5	431	107	121.39	1.13	72	1.43–48.40 (46.97)	48.51	96.83
A subgenome	1656	5208	30	5238	1686	1033.31	0.61	1418	222.41	238.69	93.72
C subgenome	1099	6250	27	6277	1126	994.22	0.88	932	310.45	406.73	76.91
Whole Genome	2755	11458	57	11515	2812	2027.53	0.72	2350	532.86	645.42	85.76

The total length of the B. napus genome was reported by Chalhoub et al. (2014; http://www.genoscope.cns.fr/brassicanapus/).

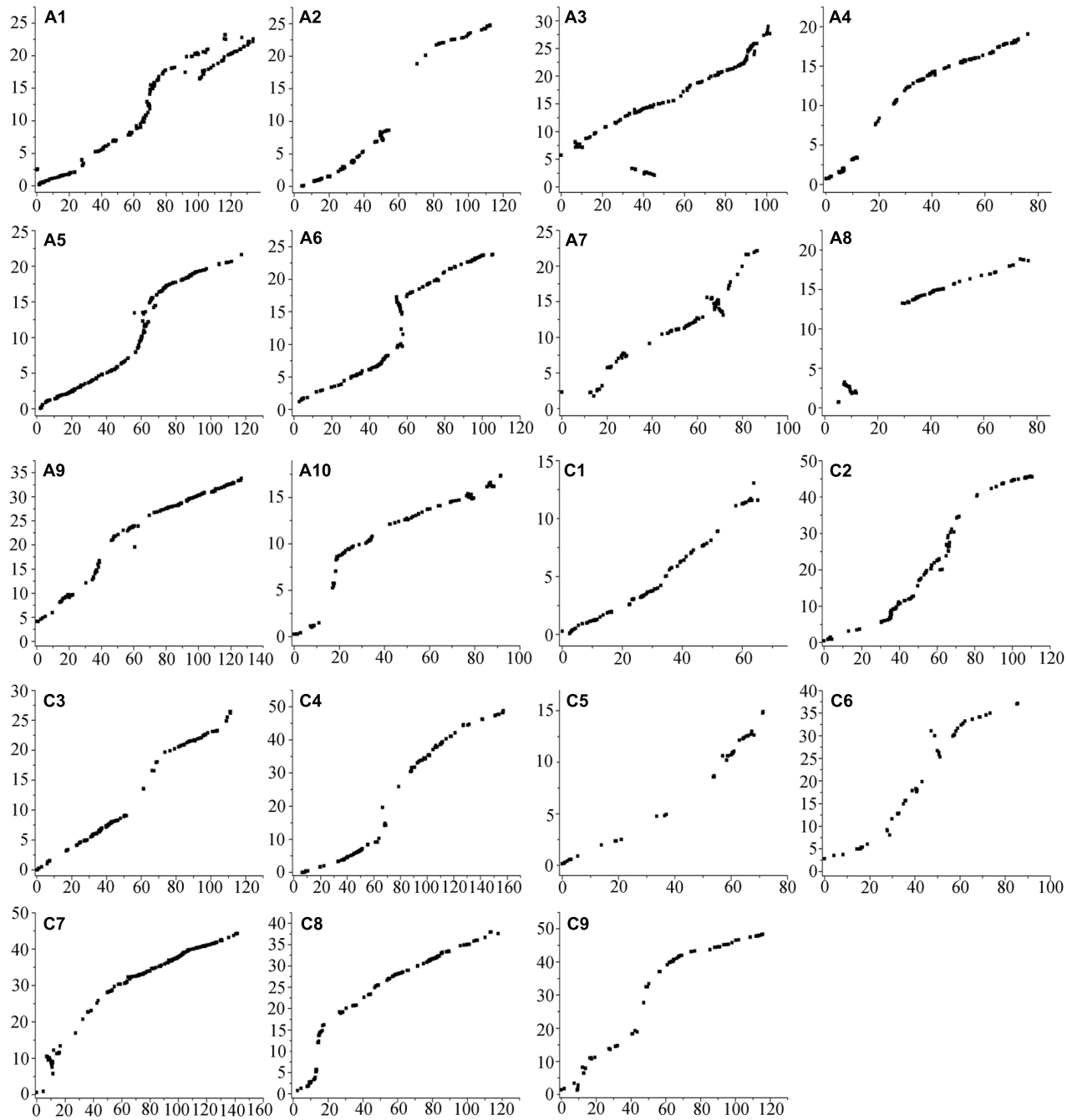


FIGURE 3 | Alignments between the SNP linkage map and the *Brassica napus* reference genome sequence. The *x*-axis represents the genetic distance in centimorgan (cM) of each linkage group in the AH map constructed in the present study. The units of the *y*-axis are the physical distances (Mb) based on the *B. napus* reference genome sequence.

for apetalous in *B. napus* have also been reported, such as the apetalous character being controlled by only one gene locus (Zhao and Wang, 2004), and four pairs of recessive genes (Lu and Fu, 1990; Zhang et al., 2007c). These studies demonstrated that PDgr in *B. napus* is a quantitative trait.

Quantitative trait loci mapping is an effective approach to dissect the genetic mechanisms of quantitative traits, while high-density map can increase the precision of QTL localization and effects, especially for small and medium sized QTLs (Almeida et al., 2013; Stange et al., 2013). However, linkage mapping studies in *B. napus* were mostly based on low-density genetic maps constructed using SSR markers, or higher-density maps based on anonymous markers, such as amplified fragment length polymorphisms or sequence-related amplified polymorphisms

TABLE 3 | A list of nine consensus QTLs for petalous degree obtained after the meta-analysis of 19 identified QTLs in five environments.

Consensus QTLs				Identified QTLs							
QTL	Peak	Marker[a]	CI	QTL	Chr.[b]	LOD	Peak	CI[c]	R^2(%)	Add[d]	Env.[e]
qPD.A3	47.82	Bn-A03-p15900189 (47.87)	46.9-49.5	*iq12NJ.A3*	A3	4.48	47.82	46.9–49.5	7.95	6.63	12NJ
qPD.A5	31.81	Bn-scaff_18338_1-p345782 (31.83)	27.6-34.0	*iq11NJ.A5*	A5	3.48	31.81	27.6–34.0	6.34	−4.51	11NJ
qPD.A6	54.8	Bn-A06-p10135140 (54.74)	49.5-55.5	*iq14DL.A6*	A6	2.82	54.8	49.5–55.5	5.08	−3.03	14DL
qPD.A9-1	64.06	Bn-A09-p25823689 (63.70)	59.66-74.36	*iq14NJ.A9*	A9	3.38	64.06	59.66–74.36	6.78	−3.69	14NJ
qPD.A9-2	76.36	Bn-A09-p29172005 (76.38)	75.97-76.75	*iq13NJ.A9*	A9	3.34	76.46	75.56–78.26	6.05	−3.95	13NJ
				iq11NJ.A9	A9	3.55	76.36	75.56–78.26	6.48	−4.52	11NJ
				iq12NJ.A9-1	A9	3.98	76.36	75.56–78.26	7.06	−4.83	12NJ
				iq14DL.A9	A9	3.26	76.36	75.56–76.46	6.12	−3.39	14DL
qPD.A9-3	81.36	Bn-A09-p30010889 (81.41)	79.86-82.96	*iq12NJ.A9-2*	A9	3.96	81.36	79.86–82.96	7.01	−4.81	12NJ
qPD.C8-1	11.8	Bn-scaff_17807_1-p364572 (11.79)	9.4-12.6	*iq11NJ.C8-1*	C8	4.63	11.8	9.4–12.6	8.77	−5.22	11NJ
qPD.C8-2	17.7	Bn-scaff_18275_1-p1278049 (17.55)	15.69-19.7	*iq11NJ.C8-2*	C8	5.59	17.6	16.7–22.8	10.47	−5.71	11NJ
				iq12NJ.C8-1	C8	6.21	17.6	14.7–25.8	11.29	−6.09	12NJ
				iq14NJ.C8	C8	3.65	17.6	15.6–26.5	7.11	−3.77	14NJ
				iq13NJ.C8-1	C8	5.48	17.9	16.7–26.5	10.24	−5.16	13NJ
				iq14DL.C8-1	C8	4.32	17.9	15.5–26.5	8.26	−3.83	14DL
qPD.C8-3	27.2	Bn-scaff_17227_1-p700248 (27.17)	26.69-27.72	*iq11NJ.C8-3*	C8	3.17	27.2	26.5–32.3	6.12	−4.42	11NJ
				iq12NJ.C8-2	C8	4.25	27.2	26.5–31.9	7.89	−5.12	12NJ
				iq13NJ.C8-2	C8	4.45	27.2	26.5–28.0	8.4	−4.7	13NJ
				iq14DL.C8-2	C8	3.69	27.2	26.5–28.0	7.06	−3.56	14DL

DL, Dali; NJ, Nanjing; 11, 12, 13 and 14 indicate the years 2011, 2012, 2013, and 2014, respectively. [a]*The closest marker and the marker position in the AH map.* [b]*Chromosome.* [c]*The 2-LOD confidence interval of QTLs.* [d]*Additive effects.* [e]*The experiment in which the QTLs were detected.*

(Sun et al., 2007). In the present study, a high-density SNP map was constructed with 2,812 loci involving 11,458 SNPs, covering a length of 2,027.53 cM and with an average marker interval of 0.72 cM, suggesting that it might be extremely useful in QTL detection. Compared with the previously published genetic maps constructed using SNPs in *B. napus* (Liu et al., 2013; Cai et al., 2014; Zhang et al., 2014), the AH map was one of the highest density maps with average distance between loci of less than 1.0 cM. In addition, the genome sequence of *B. napus* has been released (Chalhoub et al., 2014), which will facilitate the fine mapping of QTLs for quantitative traits and marker-assisted selection breeding in *B. napus*. A high-density SNP map is beneficial for the accurate alignment of the AH map to the physical chromosome segments in the assembled *B. napus* genomes. The results showed that the AH SNP map had good collinearity with the *B. napus* reference genome, indicating the high quality and accuracy of the map (**Figure 3**). Furthermore, the A sub-genome of the AH map showed a near-complete coverage of the *B. napus* genome (93.72%), considerably higher than the C sub-genome (76.91%; **Table 2**). One possible reason for the low coverage of the C sub-genome is that a form of distorted segregation in some chromosomes caused by rearrangements among homologous chromosomes in the parents occurred (Lu et al., 2013), resulting in the low coverage of C1 (33.45%), C3 (43.70%), and C5 (34.17%). The loci density in the A sub-genome (0.61 cM/marker) was also higher than that of the C sub-genome (0.88 cM/marker) in the AH map (**Table 2**), indicating the higher polymorphism rate in the A sub-genome. In agreement with these findings, Liu et al. (2013) constructed a map including 976 loci in the C genome and 1,819 loci in the A genome, with an average distance between markers of 0.53 cM in the A genome and 0.93 cM in the C genome.

A large population, a high-density genetic map and replicated experiments in multiple environments are three necessary factors for precise QTL detection (Yu et al., 2011; Wang et al., 2013). In the current study, a QTL analysis for PDgr was based on the high-density AH map and phenotypic data from five experiments, and nine consensus QTLs were identified on the A3, A5, A6, A9, and C8 chromosomes. Until now, only a few studies focused on the complex genetic mechanism of PDgr in *B. napus*. Several molecular markers associated with the petal-loss trait were located on A4 (Fray et al., 1997; Chen et al., 2006b) and C4 (Fray et al., 1997), and four QTLs for PDgr were located on A5, A6, A8, and C5 (Zhang et al., 2007a). These results suggest that the seven QTLs on A3, A9, and C8 in the present study were potential new QTLs, including the major QTL *qPD.C8-2* detected in all of the five experiments and two steadily expressed QTLs (*qPD.A9-2* and *qPD.C8-3*) identified in four experiments. The environment-specific QTLs *qPD.A5* on A5 and *qPD.A6* on A6 were not confirmed due to the lack of common markers between the different populations. The number of QTLs for PDgr in this study may be underestimated because the density of the AH genetic map was not saturated (**Table 2**), which may result in an incomplete resolution of QTLs. On the other hand, as AH population has been tested for 4 years in NJ, QTLs in DL might be underestimated based on the data of one year, and QTL detection could be improved in multiple experiments in DL location. These findings advanced our understanding of the genetic control of PDgr regulation in *B. napus*, and the

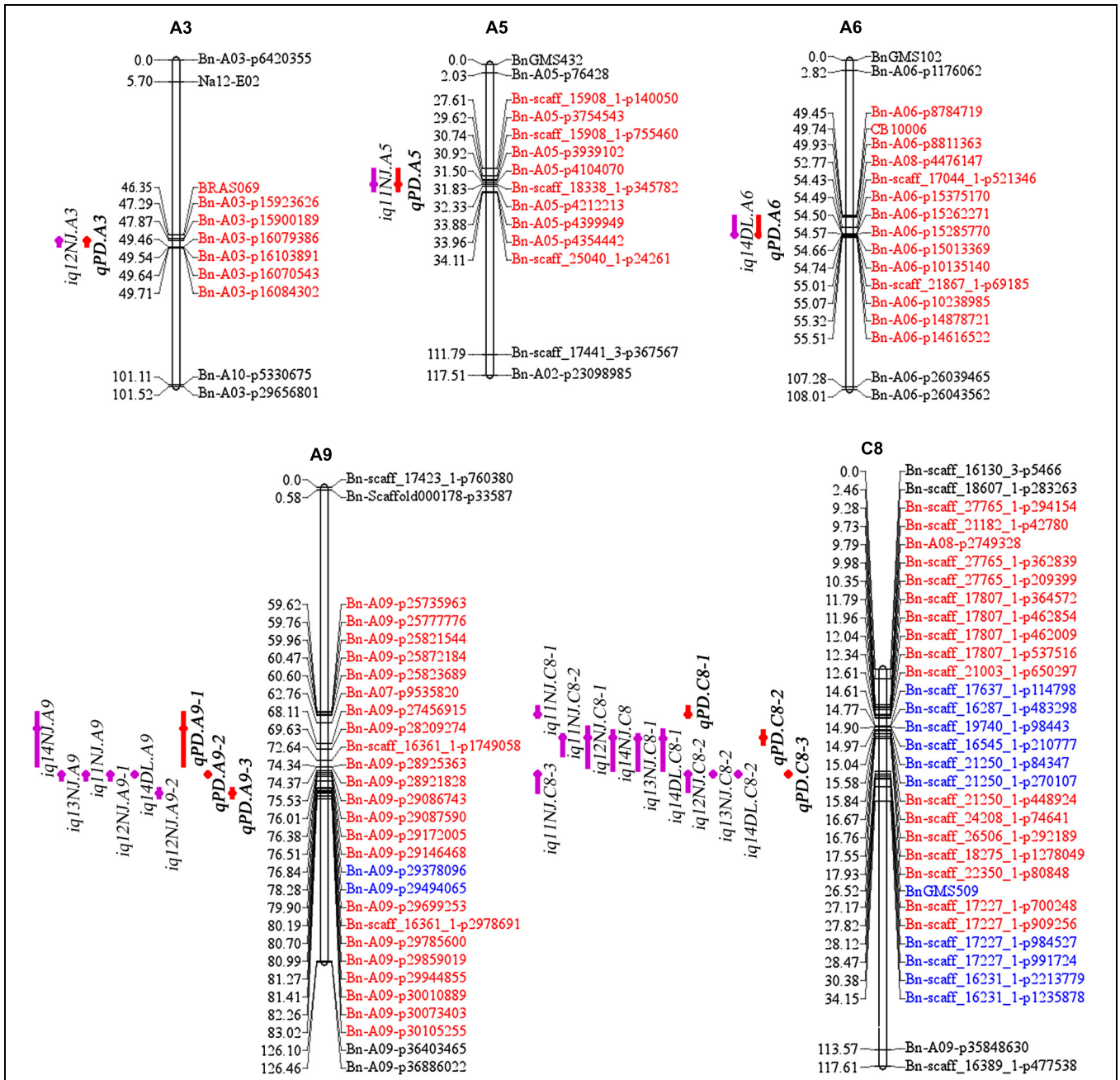


FIGURE 4 | The locations of identified and consensus QTLs associated with PDgr in the AH map. The linkage groups are represented by vertical bars. The loci names are listed on the right side of the linkage groups, while the loci positions are shown on the left side of the linkage groups. For simplicity, only the markers underlying the QTL CIs and the terminal two markers of each linkage group are shown. QTLs for PDgr are represented by bars with different backgrounds on the left side of each linkage group (purple bars, identified QTLs and red bars, consensus QTLs). Red loci are underlying the CIs of consensus QTLs and blue loci are underlying the CIs of identified QTLs but not consensus QTLs.

three stable QTLs are helpful for fine mapping and cloning of QTLs, and will enable a marker-accelerated backcrossing programs.

A typical flower consists of four different types of organs arranged in four whorls, with sepals, petals, stamens and carpels in the outermost to innermost whorls. According to the ABC model, combinatorial interactions between the three classes of floral homeotic genes are affected in the four floral organs, with 'A', 'A+B', 'B+C', and 'C' specifying sepals, petals, stamens and carpels, respectively (Coen and Meyerowitz, 1991; Honma and Goto, 2001; Theißen and Saedler, 2001). The isolation of novel floral mutants in *Arabidopsis* and other species has led to an expansion of the ABC model to include the 'D' and 'E' functions. Thus, the ABCE model was proposed, which established that the

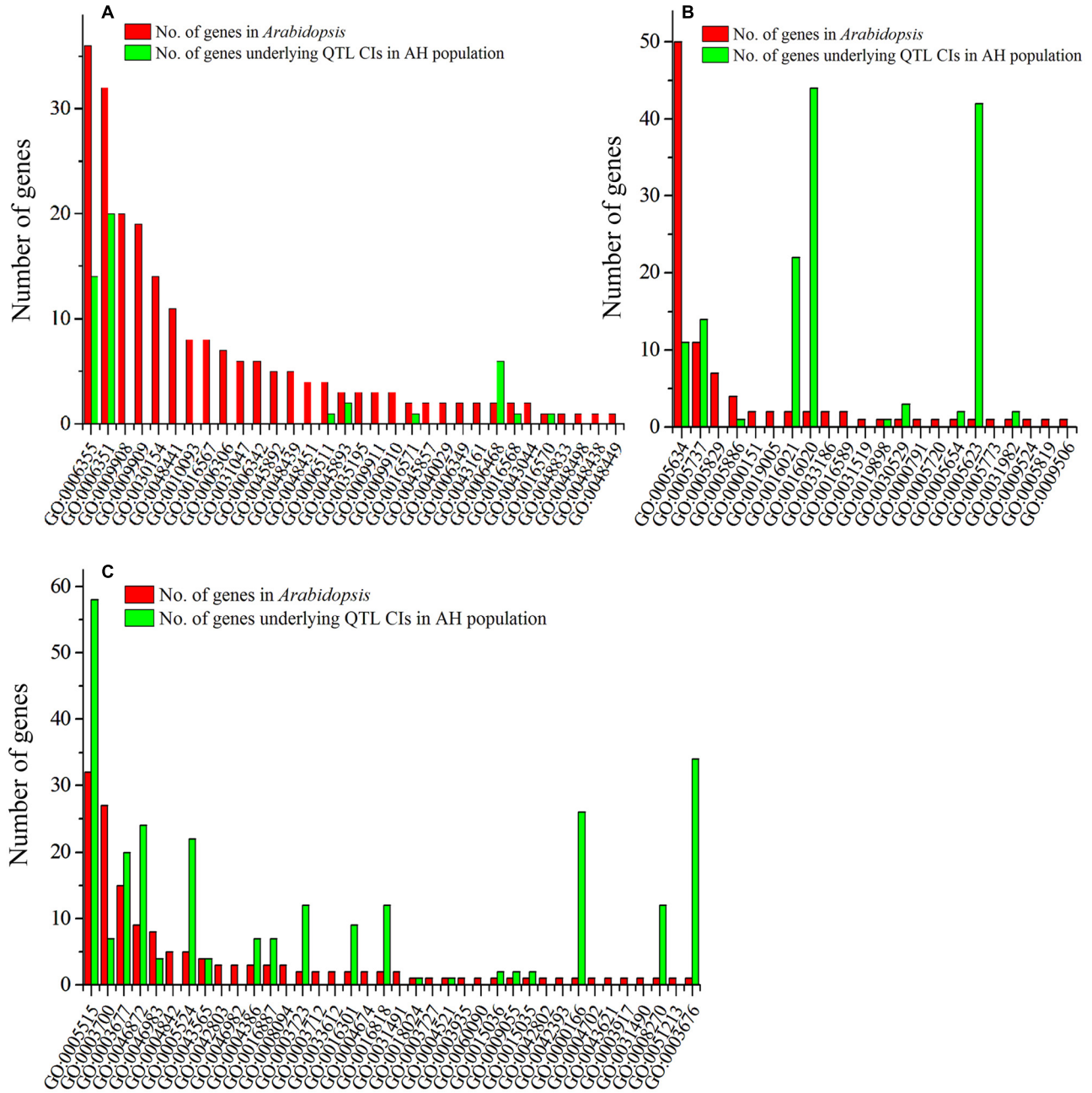


FIGURE 5 | Gene ontology (GO) annotation of collected genes regulating petal development and the candidate genes of three steady QTLs. The x-axis represents the functional groups (GO accessions), and the y-axis represents the number of genes in each category. **(A)** Biological process; **(B)** cellular component; and **(C)** molecular function.

organ-specific genes required the activity of *SEPALLATA* genes (Pinyopich et al., 2003; Krizek and Fletcher, 2005). These genes (termed class E genes), together with the class B and C genes, are required for the specification of organ identity in the petal ('A+B+E'), stamen ('B+C+E'), and carpel ('C+E') (Theißen and Saedler, 2001). Mutants with defects in the second and third whorls ('B' function) result in the homeotic conversion of petals to sepals and stamens to carpels (Causier et al., 2010). However, the apetalous line 'APL01' used in the present study had normal sepals, stamens and carpels, suggesting a more complex molecular mechanism of floral development in the allotetraploid species *B. napus*. In the present study, the genetic basis of floral development in *B. napus* was analyzed at the QTL level. Three QTLs (*qPD.A9-2*, *qPD.C8-2*, and *qPD.C8-3*) were stable across multiple environments, suggested that the genes controlling PDgr might underlying the CIs of the three QTLs.

B. napus has a common ancestor with *Arabidopsis*, and a high degree of sequence similarities and chromosomal colinearities are expected because of the progenitors diverged about 20 million years ago (Yang et al., 1999; Koch et al., 2000). Generally, the allotetraploid genomes *B. napus* may typically contain six distinct alleles for each gene present within *Arabidopsis* (Lysak et al., 2005), with the likelihood that genes carry out the core biological processes will be probable orthologs. Based on the GO assignments, 52 genes regulating of floral development mainly in *Arabidopsis* were classified into 92 functional groups, and 146 genes underlying the CIs of the three QTLs were distributed under 38 of the 92 GO terms. These genes were considered as potential candidate genes responsible for PDgr in the AH population. The B-class genes, represented in *Arabidopsis* by the MADS-box genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), experienced gene duplication events (Kramer and Hall, 2005). In *B. napus*, two types of *AP3* genes, *B.AP3.a* and *B.AP3.b*, share a high similarity in amino acid sequences, except for an eight residue difference located at the C-terminus, and, the *B.AP3.a* specified petal and stamen development and *B.AP3.b* only specified stamen development (Zhang et al., 2011). Surprisingly, no A, B, C or E class genes from the ABCE model were identified underlying the three stable QTL CIs, indicating that novel genes for PDgr in *B. napus* might exist. To gain a better understanding of how the three QTLs control PDgr in rapeseed, it is necessary to isolate these loci through a map-based cloning strategy in the future study. This study provided useful information for understanding the genetic control of floral development in *B. napus*.

AUTHOR CONTRIBUTIONS

XW and KY carried out the QTL analysis and wrote the manuscript. HL, QP, FC, and WZ participated in the field experiment. SC and MH made helpful suggestions to the manuscript. JZ designed, led, and coordinated the overall study.

ACKNOWLEDGMENTS

The work was supported by National Natural Science Foundation of China (31371660), '948' Project of Ministry of Agriculture (2011-G23), the Industry Technology System of Rapeseed in China (CARS-13), Natural Science Foundation of Jiangsu Province (BK20151369), and Jiangsu Agriculture Science and Technology Innovation Fund (CX(14)5011).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.01164

REFERENCES

Almeida, G. D., Makumbi, D., Magorokosho, C., Nair, S., Borém, A., Ribaut, J., et al. (2013). QTL mapping in three tropical maize populations reveals a set of constitutive and adaptive genomic regions for drought tolerance. *Theor. Appl. Genet.* 126, 583–600. doi: 10.1007/s00122-012-2003-7

Arcade, A., Labourdette, A., Falque, M., Mangin, B., Chardon, F., Charcosset, A., et al. (2004). BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20, 2324–2326. doi: 10.1093/bioinformatics/bth230

Basunanda, P., Radoev, M., Ecke, W., Friedt, W., Becker, H. C., and Snowdon, R. J. (2010). Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 120, 271–281. doi: 10.1007/s00122-009-1133-z

Buzza, G. (1983). The inheritance of an apetalous character in Canola (*Brassica napus*). *Cruciferae Newsletter* 8, 11–12.

Cai, G., Yang, Q., Yang, Q., Zhao, Z., Chen, H., Wu, J., et al. (2012). Identification of candidate genes of QTLs for seed weight in *Brassica napus* through comparative mapping among *Arabidopsis* and *Brassica* species. *BMC Genet.* 13:105. doi: 10.1186/1471-2156-13-105

Cai, G., Yang, Q., Yi, B., Fan, C., Edwards, D., Batley, J., et al. (2014). A complex recombination pattern in the genome of allotetraploid *Brassica napus* as revealed by a high-density genetic map. *PLoS ONE* 9:e109910. doi: 10.1371/journal.pone.0109910

Cao, X., Liu, B., and Zhang, Y. (2013). SEA: a software package of segregation analysis of quantitative traits in plants (in Chinese with an English abstract). *J. Nanjing Agric. Univ.* 36, 1–6.

Causier, B., Schwarz-Sommer, Z., and Davies, B. (2010). Floral organ identity: 20 years of ABCs. *Semin. Cell Dev. Biol.* 21, 73–79. doi: 10.1016/j.semcdb.2009.10.005

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chen, B., Wu, X., Lu, G., Gao, G., Xu, K., and Li, X. (2006a). Inheritance for apetalous trait of two kinds of apetalous rapeseed mutants (in Chinese with an English abstract). *Chin. J. Oil Crop Sci.* 28, 263–267.

Chen, B., Wu, X., Lu, G., Gao, G., Xu, K., and Li, X. (2006b). Molecular mapping of the gene (s) controlling petal-loss trait in *Brassica napus*. *Hereditas(Beijing)* 28, 707–712.

Chen, G., Geng, J., Rahman, M., Liu, X., Tu, J., Fu, T., et al. (2010). Identification of QTL for oil content, seed yield, and flowering time in oilseed rape (*Brassica napus*). *Euphytica* 175, 161–174. doi: 10.1007/s10681-010-0144-9

Chen, X., Li, X., Zhang, B., Xu, J., Wu, Z., Wang, B., et al. (2013). Detection and genotyping of restriction fragment associated polymorphisms in polyploid crops with a pseudo-reference sequence: a case study in allotetraploid *Brassica napus*. *BMC Genomics* 14:346. doi: 10.1186/1471-2164-14-346

Cheng, X., Xu, J., Xia, S., Gu, J., Yang, Y., Fu, J., et al. (2009). Development and genetic mapping of microsatellite markers from genome survey sequences in *Brassica napus*. *Theor. Appl. Genet.* 118, 1121–1131. doi: 10.1007/s00122-009-0967-8

Ching, A., Caldwell, K. S., Jung, M., Dolan, M., Smith, O. S., Tingey, S., et al. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genet.* 3:19. doi: 10.1186/1471-2156-3-19

Coen, E. S., and Meyerowitz, E. M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353, 31–37. doi: 10.1038/353031a0

Delourme, R., Falentin, C., Fomeju, B. F., Boillot, M., Lassalle, G., Andre, I., et al. (2013). High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics* 14:120. doi: 10.1186/1471-2164-14-120

Ding, G., Zhao, Z., Liao, Y., Hu, Y., Shi, L., Long, Y., et al. (2012). Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in *Brassica napus*. *Ann. Bot.* 109, 747–759. doi: 10.1093/aob/mcr323

Edwards, D., Batley, J., and Snowdon, R. J. (2013). Accessing complex crop genomes with next-generation sequencing. *Theor. Appl. Genet.* 126, 1–11. doi: 10.1007/s00122-012-1964-x

Fray, M. J., Puangsomlee, P., Goodrich, J., Coupland, G., Evans, E. J., Arthur, A. E., et al. (1997). The genetics of stamenoid petal production in oilseed rape (*Brassica napus*) and equivalent variation in *Arabidopsis thaliana*. *Theor. Appl. Genet.* 94, 731–736. doi: 10.1007/s001220050472

Habekotté, B. (1997). Options for increasing seed yield of winter oilseed rape (*Brassica napus* L.) a simulation study. *Field Crop Res.* 54, 109–126. doi: 10.1016/S0378-4290(97)00041-5

Hoekman, B. (2007). *Global Economic Prospects 2008: Technology Diffusion in the Developing World*. Washington, DC: World Bank Publications, 40–41.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409, 525–529. doi: 10.1038/35054083

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695

Jamaux, D. I., and Spire, D. (1999). Comparison of responses of ascospores and mycelium by ELISA with anti-mycelium and anti-ascospore antisera for the development of a method to detect *Sclerotinia sclerotiorum* on petals of oilseed rape. *Ann. Appl. Biol.* 134, 171–179. doi: 10.1111/j.1744-7348.1999.tb05253.x

Jiang, L. (2007). Breeding for apetalous rape: inheritance and yield physiology. *Adv. Bot. Res.* 45, 217–231. doi: 10.1016/S0065-2296(07)45008-X

Jiang, L., and Becker, H. (2003). Inheritance of apetalous flowers in a mutant of oilseed rape. *Crop Sci.* 43:508. doi: 10.2135/cropsci2003.0508

Kelly, A., Fray, M., Arthur, E. A., Lydiate, D. J., and Evans, E. J. (1995). "The genetic control of petalless flowers and upright pods," in *Proceedings of the 9th International Rapeseed Congress*, Cambridge, 4–8.

Koch, M. A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (*Brassica*ceae). *Mol. Biol. Evol.* 17, 1483–1498. doi: 10.1093/oxfordjournals.molbev.a026248

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Kramer, E. M., and Hall, J. C. (2005). Evolutionary dynamics of genes controlling floral development. *Curr. Opin. Plant. Biol.* 8, 13–18. doi: 10.1016/j.pbi.2004.09.019

Krizek, B. A., and Fletcher, J. C. (2005). Molecular mechanisms of flower development: an armchair guide. *Nat. Rev. Genet.* 6, 688–698. doi: 10.1038/nrg1675

Landry, B. S., Hubert, N., Etoh, T., Harada, J. J., and Lincoln, S. E. (1991). A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. *Genome* 34, 543–552. doi: 10.1139/g91-084

Lefol, C., and Morrall, R. A. A. (1996). Immunofluorescent staining of *Sclerotinia* ascospores on canola petals. *Can. J. Plant Pathol.* 18, 237–241. doi: 10.1080/07060669609500618

Li, H., Yu, K., Guo, T., and Zhang, J. (2014). Genetic analysis of apetalous trait in *Brassica napus* using mixed model of major gene and polygene (in Chinese with an English abstract). *Jiangsu J. Agric. Sci.* 30, 253–258.

Linnemann, A. R., and Dijkstra, D. S. (2002). Toward sustainable production of protein-rich foods: appraisal of eight crops for Western Europe. Part I. Analysis of the primary links of the production chain. *Crit. Rev. Food Sci. Nutr.* 42, 377–401. doi: 10.1080/20024091054193

Liu, L., Qu, C., Wittkop, B., Yi, B., Xiao, Y., He, Y., et al. (2013). A high-density SNP map for accurate mapping of seed fibre QTL in *Brassica napus* L. *PLoS ONE* 8:e83052. doi: 10.1371/journal.pone.0083052

Long, Y., Shi, J., Qiu, D., Li, R., Zhang, C., Wang, J., et al. (2007). Flowering time quantitative trait loci analysis of oilseed *Brassica* in multiple environments and genomewide alignment with *Arabidopsis*. *Genetics* 177, 2433–2444. doi: 10.1534/genetics.107.080705

Lowe, A. J., Moule, C., Trick, M., and Edwards, K. J. (2004). Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. *Theor. Appl. Genet.* 108, 1103–1112. doi: 10.1007/s00122-003-1522-7

Lu, W., Liu, J., Xin, Q., Wan, L., Hong, D., and Yang, G. (2013). A triallelic genetic male sterility locus in *Brassica napus*: an integrative strategy for its physical mapping and possible local chromosome evolution around it. *Ann. Bot.* 111, 305–315. doi: 10.1093/aob/mcs260

Lu, Z., and Fu, S. (1990). Inheritance of apetalous character in rape (*Brassica napus* L.) and its implication in breeding (in Chinese with an English abstract). *Jiangsu J. Agric. Sci.* 6, 30–36.

Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. *Genome Res.* 15, 516–525. doi: 10.1101/gr.3531105

Mauricio, R. (2001). Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370–381. doi: 10.1038/35072085

McCouch, S. R., Cho, Y. G., Yano, M., Paul, E., Blinstrub, M., Morishima, H., et al. (1997). Report on QTL nomenclature. *Rice Genet. News* 14, 11–131. doi: 10.1007/s10142-013-0328-1

Mendham, N. J., Rao, M. S. S., and Buzza, G. C. (1991). "The apetalous flower character as a component of a high yielding ideotype," in *Proceedings of the 8th International Rapeseed Congress*, Saskatoon, SK, 596–600.

Mendham, N. J., Shipway, P. A., and Scott, R. K. (1981). The effects of delayed sowing and weather on growth, development and yield of winter oil-seed rape (*Brassica napus*). *J. Agric. Sci.* 96, 389–416. doi: 10.1017/S002185960006617X

Mochida, K., Yamazaki, Y., and Ogihara, Y. (2003). Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. *Mol. Genet. Genomics* 270, 371–377. doi: 10.1007/s00438-003-0939-7

Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., et al. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature* 424, 85–88. doi: 10.1038/nature01741

Piquemal, J., Cinquin, E., Couton, F., Rondeau, C., Seignoret, E., Doucet, I., et al. (2005). Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theor. Appl. Genet.* 111, 1514–1523. doi: 10.1007/s00122-005-0080-6

Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., et al. (2006). A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theor. Appl. Genet.* 114, 67–80. doi: 10.1007/s00122-006-0411-2

Quijada, P. A., Udall, J. A., Lambert, B., and Osborn, T. C. (2006). Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.)*:* 1. Identification of genomic regions from winter germplasm. *Theor. Appl. Genet.* 113, 549–561. doi: 10.1007/s00122-006-0323-1

Raman, H., Dalton-Morgan, J., Diffey, S., Raman, R., Alamery, S., Edwards, D., et al. (2014). SNP markers-based map construction and genome-wide linkage analysis in *Brassica napus*. *Plant Biotechnol. J.* 12, 851–860. doi: 10.1111/pbi.12186

Ramanujam, S. (1940). An apetalous mutation in turnip (*Brassica* campestris L.). *Nature* 145, 552–553. doi: 10.1038/145552a0

Shi, J., Li, R., Qiu, D., Jiang, C., Long, Y., Morgan, C., et al. (2009). Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. *Genetics* 182, 851–861. doi: 10.1534/genetics.109.101642

Snowdon, R. J., and Iniguez Luy, F. L. (2012). Potential to improve oilseed rape and canola breeding in the genomics era. *Plant Breed.* 131, 351–360. doi: 10.1111/j.1439-0523.2012.01976.x

Stange, M., Utz, H. F., Schrag, T. A., Melchinger, A. E., and Würschum, T. (2013). High-density genotyping: an overkill for QTL mapping? Lessons learned from a case study in maize and simulations. *Theor. Appl. Genet.* 126, 2563–2574. doi: 10.1007/s00122-013-2155-0

Sun, Z., Wang, Z., Tu, J., Zhang, J., Yu, F., McVetty, P. B. E., et al. (2007). An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theor. Appl. Genet.* 114, 1305–1317. doi: 10.1007/s00122-006-0483-z

Tan, X., Li, D., Wang, H., and Guo, A. (2003). Identification of a RAPD marker linked to a petal-controlled gene in *Brassica napus* (in Chinese with an English abstract). *Acta Bot. Boreal. Occident. Sin.* 23, 561–565.

Theißen, G., and Saedler, H. (2001). Plant biology: floral quartets. *Nature* 409, 469–471. doi: 10.1038/35054172

Udall, J. A., Quijada, P. A., Lambert, B., and Osborn, T. C. (2006). Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.)*:* 2. Identification of alleles from unadapted germplasm. *Theor. Appl. Genet.* 113, 597–609. doi: 10.1007/s00122-006-0324-0

Uzunova, M. I., and Ecke, W. (1999). Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). *Plant Breed.* 118, 323–326. doi: 10.1046/j.1439-0523.1999.00371.x

Van Ooijen, J. W. (2006). *JoinMap 4. Software for the Calculation of Genetic Linkage Maps in Experimental Populations*. Kyazma BV.

Vignal, A., Milan, D., SanCristobal, M., and Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genet. Sel. Evol.* 34, 275–305. doi: 10.1051/gse:2002009

Virk, P. S., Khush, G. S., and Peng, S. (2004). Breeding to enhance yield potential of rice at IRRI: the ideotype approach. *Int. Rice Res. Notes* 29, 5–9.

Wang, S., Basten, C. J., and Zeng, Z. B. (2007). *Windows QTL Cartographer 2.5*. Raleigh, NC: North Carolina State University.

Wang, X., Long, Y., Yin, Y., Zhang, C., Gan, L., Liu, L., et al. (2015a). New insights into the genetic networks affecting seed fatty acid concentrations in *Brassica napus*. *BMC Plant Biol.* 15:91. doi: 10.1186/s12870-015-0475-8

Wang, X., Wang, H., Long, Y., Liu, L., Zhao, Y., Tian, J., et al. (2015b). Dynamic and comparative QTL analysis for plant height in different developmental stages of *Brassica napus* L. *Theor. Appl. Genet.* 128, 1175–1192. doi: 10.1007/s00122-015-2498-9

Wang, X., Wang, H., Long, Y., Li, D., Yin, Y., Tian, J., et al. (2013). Identification of QTLs associated with oil content in a high-oil *Brassica napus* cultivar and construction of a high-density consensus map for QTLs comparison in *B. napus*. *PLoS ONE* 8:e80569. doi: 10.1371/journal.pone.0080569

Yang, Y., Lai, K., Tai, P., and Li, W. (1999). Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J. Mol. Evol.* 48, 597–604. doi: 10.1007/PL00006502

Yu, H., Xie, W., Wang, J., Xing, Y., Xu, C., Li, X., et al. (2011). Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS ONE* 6:e17595. doi: 10.1371/journal.pone.0017595

Zeng, Z. (1994). Precision mapping of quantitative trait loci. *Genetics* 136, 1457–1468.

Zhang, D., Hua, Y., Wang, X., Zhao, H., Shi, L., and Xu, F. (2014). A high-density genetic map identifies a novel major QTL for boron efficiency in oilseed rape (*Brassica napus* L.). *PLoS ONE* 9:e112089. doi: 10.1371/journal.pone.0112089

Zhang, J., Fu, S., Qi, C., Pu, H., and Wu, X. (2002). Preliminary report of breeding for near-isogenic lines of apetalous in rapeseed (*B. napus* L.) (in Chinese with an English abstract). *Chin. J. Oil Crop Sci.* 24, 14–16.

Zhang, J., Qi, C., Li, G., Pu, H., Chen, S., Chen, X., et al. (2007a). Genetic map construction and apetalousness QTLs identification in rapeseed (*Brassica napus* L.) (in Chinese with an English abstract). *Acta. Agron. Sin.* 33, 1246–1254.

Zhang, J., Qi, C., Pu, H., Chen, S., Chen, F., Chen, X., et al. (2007b). Genetic analysis of apetalous in *Brassica napus* L. using mixed model of major gene and polygene (in Chinese with an English abstract). *Chin. J. Oil Crop Sci.* 29, 227–232.

Zhang, R., Tang, X., Li, M., and Chen, D. (2007c). Study on genetics of apetalous line NF001 in rapeseed (*B. napus* L.) (in Chinese with an English abstract). *Seed* 26, 13–15.

Zhang, Y., Wang, X., Zhang, W., Yu, F., Tian, J., Li, D., et al. (2011). Functional analysis of the two *Brassica* AP3 genes involved in apetalous and stamen carpelloid phenotypes. *PLoS ONE* 6:e20930. doi: 10.1371/journal.pone.0020930

Zhao, Y., and Wang, M. (2004). Inheritance and agronomic performance of an apetalous flower mutant in *Brassica napus* L. *Euphytica* 137, 381–386. doi: 10.1023/B:EUPH.0000040522.37048.7a

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Wang, Yu, Li, Peng, Chen, Zhang, Chen, Hu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 January 2016
doi: 10.3389/fpls.2016.00017

Genome-Wide Identification of QTL for Seed Yield and Yield-Related Traits and Construction of a High-Density Consensus Map for QTL Comparison in *Brassica napus*

Weiguo Zhao[1,2†], Xiaodong Wang[3†], Hao Wang[1,2], Jianhua Tian[1], Baojun Li[1], Li Chen[2], Hongbo Chao[2], Yan Long[4], Jun Xiang[5], Jianping Gan[5], Wusheng Liang[6] and Maoteng Li[2,5*]*

[1] Hybrid Rapeseed Research Center of Shaanxi Province, Shaanxi Rapeseed Branch of National Centre for Oil Crops Genetic Improvement, Yangling, China, [2] Department of Biotechnology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, China, [3] Key Laboratory of Cotton and Rapeseed, Ministry of Agriculture, Institute of Industrial Crops, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [4] Institute of Biotechnology, Chinese Academy of Agricultural Sciences, Beijing, China, [5] Hubei Collaborative Innovation Center for the Characteristic Resources Exploitation of Dabie Mountains, Huanggang Normal University, Huanggang, China, [6] Department of Applied Biological Science, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China

OPEN ACCESS

Edited by:
Juan Francisco Jimenez Bremont, Instituto Potosino de Investigacion Cientifica y Tecnologica, Mexico

Reviewed by:
Yinglong Chen, The University of Western Australia, Australia
Daniela Marone, Cereal Research Centre - CRA-CER, Italy

***Correspondence:**
Hao Wang
wangzy846@sohu.com;
Maoteng Li
limaoteng426@mail.hust.edu.cn

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *25 August 2015*
Accepted: *08 January 2016*
Published: *28 January 2016*

Citation:
Zhao W, Wang X, Wang H, Tian J, Li B, Chen L, Chao H, Long Y, Xiang J, Gan J, Liang W and Li M (2016) Genome-Wide Identification of QTL for Seed Yield and Yield-Related Traits and Construction of a High-Density Consensus Map for QTL Comparison in Brassica napus. Front. Plant Sci. 7:17. doi: 10.3389/fpls.2016.00017

Seed yield (SY) is the most important trait in rapeseed, is determined by multiple seed yield-related traits (SYRTs) and is also easily subject to environmental influence. Many quantitative trait loci (QTLs) for SY and SYRTs have been reported in *Brassica napus*; however, no studies have focused on seven agronomic traits simultaneously affecting SY. Genome-wide QTL analysis for SY and seven SYRTs in eight environments was conducted in a doubled haploid population containing 348 lines. Totally, 18 and 208 QTLs for SY and SYRTs were observed, respectively, and then these QTLs were integrated into 144 consensus QTLs using a meta-analysis. Three major QTLs for SY were observed, including *cqSY-C6-2* and *cqSY-C6-3* that were expressed stably in winter cultivation area for 3 years and *cqSY-A2-2* only expressed in spring rapeseed area. Trait-by-trait meta-analysis revealed that the 144 consensus QTLs were integrated into 72 pleiotropic unique QTLs. Among them, all the unique QTLs affected SY, except for *uq.A6-1*, including *uq.A2-3*, *uq.C1-2*, *uq.C1-3*, *uq.C6-1*, *uq.C6-5,* and *uq.C6-6* could also affect more than two SYRTs. According to the constructed high-density consensus map and QTL comparison from literatures, 36 QTLs from five populations were co-localized with QTLs identified in this study. In addition, 13 orthologous genes were observed, including five each gene for SY and thousand seed weight, and one gene each for biomass yield, branch height, and plant height. The genomic information of these QTLs will be valuable in hybrid cultivar breeding and in analyzing QTL expression in different environments.

Keywords: *Brassica napus*, seed yield, seed yield-related traits, quantitative trait loci, map comparsion, candidate genes

Abbreviations: *B. napus, Brassica napus; B. rapa, Brassica rapa; B. oleracea, Brassica oleracea;* QTL, Quantitative trait locus; DH, Doubled haploid; cM, CentiMorgan; CIs, Confidence intervals; PVE, Phenotypic variation explained; SY, Seed yield; SYRTs, Seed yield related traits; BY, Biomass yield; SW, Thousand seed weight; BH, First effective branch height; PH, Plant height; FBN, First effective branch number; LMI, Length of main inflorescence; PMI, Pod number of main inflorescence; LOD, Log odds score methods.

INTRODUCTION

Brassica napus (AACC, $2n = 38$) originated from hybridization between *Brassica rapa* (AA, $2n = 20$) and *Brassica oleracea* (CC, $2n = 18$; UN, 1935), and is the second most important oilseed crop after soybean (Basunanda et al., 2010). As the global requirements for rapeseed oil and protein are growing rapidly, increasing seed yield (SY) is the main breeding aim at present. SY is directly determined by yield component traits, including thousand seed weight (SW), pod number per plant and seed number per pod (Qzer et al., 1999; Quarrie et al., 2006). In addition, SY is also indirectly influenced by other seed yield related traits (SYRTs), such as biomass yield (BY), plant height (PH), first effective branch height (BH), first effective branch number (FBN), length of main inflorescence (LMI), and pod number of main inflorescence (PMI) in *B. napus* (Qiu et al., 2006; Li et al., 2007; Shi et al., 2009). Interactions between SY, SW, PH, BH, FBN, LMI, and PMI were observed in previous studies (Yu, 1998; Zhang et al., 2006).

SY and SYRTs are all complex quantitative traits controlled by multiple genes (Kearsey and Pooni, 1998). QTL analysis has proved a powerful genetic approach to dissect complex traits (Paran and Zamir, 2003). Many QTLs for SY and SYRTs have been reported in *B. napus*, such as QTLs for SY being mainly located on A10, C3, and C6 (Quijada et al., 2006; Udall et al., 2006; Maccaferri et al., 2008). In addition, studies related to QTLs for SY and/or several SYRTs have also been performed (Chen et al., 2007, 2010; Li et al., 2007; Maccaferri et al., 2008; Shi et al., 2009; Basunanda et al., 2010; Ding et al., 2012; Cai et al., 2014). As the genetic backgrounds of different mapping populations for *B. napus* vary considerably, the number and location of QTLs detected in different populations also differ, thus is very necessary to contrast the QTLs for SY and SYRTs and select the common QTLs in different populations. Although many QTLs for SY and SYRTs have been reported, studies that simultaneously focused on the eight agronomic traits (SY, BY, SW, PH, BH, FBN, LMI, and PMI) are rare. Moreover, the candidate genes for these QTLs have rarely been mentioned. Comparative mapping among the model plant *Arabidopsis thaliana* with related species is a powerful tool to identify candidate genes. For example, Long et al. (2007) obtained the candidate gene *BnFLC10* underlying QTL *qFT10-4* and identified the key gene controlling differentiation of winter or spring type rapeseed based on comparative mapping analysis. Shi et al. (2009) and Ding et al. (2011) also obtained the candidate genes controlling flower time and seed phosphorus concentration, respectively, by comparative mapping with the *Arabidopsis* genome. Comparative mapping among *B. napus*, *Arabidopsis*, *B. rapa*, and *B. oleracea* genomes is necessary to obtain candidate genes in the confidence intervals (CIs) of QTLs for SY and SYRTs.

In order to increase statistical power and precision of obtaining QTLs, a high-density genetic linkage map is considered as a key factor (Jiang and Zeng, 1995). Several high-density genetic maps for *B. napus* have been constructed by integrating different linkage maps based on common molecular markers from different populations (Lombard and Delourme, 2001; Scoles et al., 2007; Raman et al., 2013). For example, Lombard and Delourme (2001) constructed a consensus map covering a total length of 2429.0 cM by integrating three individual linkage maps, and Wang et al. (2013) constructed a high-density consensus map with 1335 markers covering 2395.2 cM of the total genome length by merging eight individual linkage maps from different populations. Zhou et al. (2014) used 15 published articles concerning *B. napus* mapping experiments over the last decade and carried out *in silico* integration of 1960 QTLs with 13 SY and SYRTs, a total of 736 QTLs were mapped onto 283 loci in the A and C genomes of *B. napus*.

In the present study, a large doubled haploid (DH) population containing 348 lines was used to investigate the QTLs for SY and SYRTs in multiple environments, and then a consensus map was constructed for QTL comparison between the KN (the population used in this study) and five other published populations. These results provide abundant useful information to further understanding of the genetic mechanisms of SY and SYRTs, and could be used in marker assisted selection for improving SY in *B. napus*.

MATERIALS AND METHODS

Plant Material and Field Experiments

A DH population, named KN and containing 348 lines derived from KenC-8 and N53-2, was used in this study (Wang et al., 2013). The KN genetic linkage map was constructed with 403 molecular markers, including 275 simple sequence repeats, 117 sequence-related amplified polymorphisms, 10 sequence tagged sites, and one intron fragment length polymorphism, which covered a total length of 1783.9 cM. The KN population and its parents were grown in eight environments, including a winter rapeseed area, Dali of Shannxi Province (coded DL), in northwest China for five successive years (September–May of 2008–2009, 2009–2010, 2010–2011, 2011–2012, and 2012–2013); a spring rapeseed area, Sunan of Gansu Province (coded GS), in northwest China for three successive years (April–September of 2010, 2011, and 2012). Year-location combinations were treated as micro-environments, for example, 09DL means that the experiment was carried out in September–May of 2009–2010 at DL. Meanwhile, each year-location combination was treated as a trial.

The field experiments followed a randomized complete block design. The KN population, together with the two parents, was planted in DL and GS with three and two replications, respectively. Each field trial consisted of 348 lines. Each line was grown in a two-row plot with 40 cm between rows and 20 cm between individuals, and row length of 250 cm in all environments.

Measurement of Phenotypic Data of SY and SYRTs

Phenotypic data for SY (g/plant) were recorded with five representative plants in the middle of each plot. These five plants were also used for measurement of other SYRTs: BY (g/plant), SW (g), PH (cm), BH (cm), FBN, LMI (cm), and PMI. Because of a strong requirement for vernalization, the N53-2 and some DH lines did not flower or fully mature in the spring area (10GS, 11GS, and 12GS), and so SY of these DH lines were treated as

missing data. SY was the average dry weight of seeds of the five representative individuals. BY was measured as the average total above-ground dry weight of the five plants (excluding the seeds). SW was the average dry weight of 1000 well-filled seeds from the five samples. PH was the average height of the five individuals, measured from the base of the stem to the tip of the main inflorescence. FBN was the number of branches arising from the main stem of each harvested individual. LMI was measured from the bottom to the top of the main inflorescence. PMI was effective pod number from the bottom to the top of the main inflorescence.

Statistical Analysis, QTL Mapping and Meta-Analysis

Estimates of means and variances for the SY and SYRTs were implemented using SPSS 18.0 software (SPSS Inc., Chicago, IL, USA). The QTL information for PH was according to Wang et al. (2015) in 10DL, 11DL, 12DL, 11GS, 12GS, and 12WH. QTL detection for other traits was conducted by composite interval mapping with Windows QTL Cartographer 2.5 software (Wang et al., 2007). The estimated additive effect and phenotypic variation explained (PVE) by each putative QTL were obtained using composite interval mapping model. Significance levels for the log odds score methods (LOD) were determined by 1000-permutation test corresponding to $P = 0.05$, and LOD of 2.8–3.1 was used to, respectively, identify significant QTLs in each environment, and these QTLs were termed "identified QTL." QTLs that mapped to the same region with overlapping CIs were assumed to be the same, and BioMercator 2.1 software was used to integrate these QTLs into consensus QTLs using the meta-analysis method (Arcade et al., 2004). If a consensus QTL had at least one environment with PVE $\geq$ 20% or at least two environments with PVE $\geq$ 10%, the QTL was defined as a major QTL; the remaining QTLs were defined as minor QTLs (Maccaferri et al., 2008).

The identified QTLs for SY and SYRTs were named according to Wang et al. (2013); for example, the QTL abbreviation "*qSY*" (*q*, QTL; *SY*, seed yield) suffixed with the linkage group (A1–A10, C1–C9), a hyphen (-), and finally the serial number of QTLs in the linkage group (e.g., *qSY-A2-1*). The QTL integrations were adopted by meta-analysis; for example, the identified QTLs were integrated into consensus QTLs trait-by-trait, and the consensus QTLs for SY and SYRTs with overlapping CIs were integrated into pleiotropic unique QTLs using BioMercator 2.1 (Arcade et al., 2004). The name of consensus QTLs and unique QTLs referred to the name of identified QTLs. For each unique QTL, one or more consensus QTLs for SYRTs were chosen as indicator QTLs, which were defined as potential genetic determinants of the co-localized QTL for SY.

QTL Projection from Other Populations onto the KN Map, and QTL Comparison among the Different Populations

The map projection package of BioMercator 2.1 software was used for QTL projection of SY and SYRTs in five previously reported populations onto the KN genetic map (**Table 1**), including the QN (Quantum × No. 2127-17; Chen et al., 2007), SE (SI-1300 × Eagle; Li et al., 2007), ER (Express617 × R53; Radoev et al., 2008), TN (Tapidor × Ningyou7; Shi et al., 2009), and BE (B104-2 × Eyou Changjia; Ding et al., 2012) populations. The method for projection of QTLs from different linkage groups was according to Arcade et al. (2004). The method for QTL comparison from different linkage groups was a "two-round" strategy (Shi et al., 2013), and the detailed methods for QTL comparison from different populations were found in Ding et al. (2012) and Jiang et al. (2014). The consensus QTLs from different populations were named with the population abbreviation followed by the consensus QTL names for QTL comparison (e.g., *KNcqSY-A2-1*).

Candidate Gene Observations by Comparative Mapping among *Arabidopsis, B. rapa, B. oleracea, and B. napus*

Among the 403 markers mapped in the KN genetic map (Table S1), 141 markers with known sequence information were used for the sequence comparisons of the *Arabidopsis* genome database with other *Brassica species* (http://www.arabidopsis.org/). The databases of *B. oleracea* (Liu et al., 2014), *B. rapa* (http://brassicadb.org/brad/), and *B. napus* (http://www.genoscope.fr/brassicanapus/) were used for confirmation of homologous genes on the genomes of *Arabidopsis*, *B. rapa*, *B. oleracea,* and *B. napus.* Firstly, the 141 markers with known sequence information were used as anchored markers to carry out map alignment between *B. napus* and *Arabidopsis* according to the method of Long et al. (2007). If three or more sequence informative markers in the KN population were closely linked within one conserved block of *Arabidopsis* (Schranz et al., 2006), a synteny block was considered to exist. If there were only one or two sequence informative marker(s), this was recognized as an insertion segment. Secondly, if a synteny block or insertion segment were co-localized with the CI of a QTL, the genes underlying the synteny block or

TABLE 1 | Five reported populations were used for QTL projection of SY and SYRTs in *B. napus*.

Population parental lines	Abbreviation	Size	Type	Traits	References
Tapidor × Ningyou7	TN	202	DH	BN BY PH SW SY	Shi et al., 2009
SI-1300 × Eagle	SE	184	F_2	PH HPB ELMI SMI	Li et al., 2007
B104-2 × Eyou Changjia	BE	124	F_2	SY SW BN PH	Ding et al., 2012
Express617 × R53	ER	250	DH	SY SW	Radoev et al., 2008
Quantum × No. 2127-17	QN	258	DH	PH HPB LMI FB	Chen et al., 2007

TABLE 2 | Mean values and phenotypic variation for SY and SYRTs of KN population evaluated in different micro-environments.

SY(g)[a]		08DL[c]	09DL	10DL	10GS	11DL	11GS	12DL	12GS
KenC-8	Mean[b]	20.08 ± 1.82	24.65 ± 4.81	24.25 ± 6.80	24.68 ± 1.22	16.22 ± 0.91	22.37 ± 1.73	18.77 ± 2.03	23.23 ± 3.63
N53-2	Mean	27.00 ± 4.59	24.23 ± 4.91	25.09 ± 4.41		25.29 ± 2.86		29.18 ± 3.50	
DH	Mean	19.48 ± 5.38	22.29 ± 5.13	22.87 ± 5.74	18.47 ± 5.59	18.76 ± 5.46	16.91 ± 5.15	23.83 ± 5.45	14.36 ± 4.25
Range	Min-Max	2.24–34.91	5.96–35.07	4.3–38.82	8.1–35.1	2.45–34.81	6.43–33.74	6.04–40.83	4.12–26.37
BY(g)		**08DL**	**09DL**	**10DL**	**10GS**	**11DL**	**11GS**	**12DL**	**12GS**
KenC-8	Mean	58.27 ± 4.49	74.49 ± 20.09	70.69 ± 18.56	106.63 ± 14.95	59.99 ± 5.31	91.66 ± 15.33	39.49 ± 5.11	57.09 ± 1.60
N53-2	Mean	86.74 ± 13.92	71.04 ± 14.84	76.92 ± 12.52		71.37 ± 6.75		45.81 ± 8.43	
DH	Mean	62.68 ± 17.14	46.44 ± 9.83	71.08 ± 16.04	84.49 ± 19.28	59.52 ± 14.78	53.03 ± 11.56	45.61 ± 9.62	48.08 ± 13.54
Range	Min–Max	9.7–107.1	17.1–74.0	16.4–108.2	41.8–135.7	12.9–97.9	27.9–95.0	14.0–116.6	14.0–116.6
SW(g)		**08DL**	**09DL**	**10DL**	**10GS**	**11DL**	**11GS**	**12DL**	**12GS**
KenC-8	Mean	3.53 ± 0.31	3.72 ± 0.31	3.34 ± 0.43	3.29 ± 0.01	2.70 ± 0.50	3.15 ± 0.54	3.00 ± 0.49	3.15 ± 0.54
N53-2	Mean	3.93 ± 0.10	4.02 ± 0.29	3.47 ± 0.47		2.88 ± 0.23		3.33 ± 0.12	
DH	Mean	3.43 ± 0.55	3.76 ± 0.47	3.31 ± 0.51	3.15 ± 0.52	2.81 ± 0.52	3.45 ± 0.56	3.3 ± 0.46	2.92 ± 0.72
Range	Min–Max	1.48–4.83	2.47–4.92	1.91–5.71	1.73–5.20	1.57–5.78	2.00–5.93	2.24–5.64	1.11–5.24
PH(cm)		**08DL**	**09DL**	**10DL**	**10GS**	**11DL**	**11GS**	**12DL**	**12GS**
KenC-8	Mean	120.00 ± 10.00	144.00 ± 6.56	145.67 ± 6.03	172.45 ± 7.84	151.67 ± 11.59	191.00 ± 1.41	135.67 ± 9.07	208.50 ± 6.36
N53-2	Mean	145.00 ± 13.23	156.33 ± 13.43	166.00 ± 5.57	176.87 ± 10.78	165.33 ± 11.93	181.50 ± 27.58	156.67 ± 11.06	228.00 ± 7.07
DH	Mean	128.29 ± 16.64	148.66 ± 16.97	143.28 ± 18.49	185.18 ± 24.86	152.52 ± 19.36	189.31 ± 30.44	132.87 ± 17.64	196.95 ± 25.03
Range	Min–Max	83.3–163.3	90.0–175.7	94.3–182.0	108.0–249.1	94.7–193.0	132.0–296.5	87.3–170.3	103.0–261.2
BH(cm)		**08DL**	**09DL**	**10DL**	**11DL**	**11GS**	**12DL**	**12GS**	
KenC-8	Mean	27.04 ± 2.25	47.47 ± 5.01	36.20 ± 10.11	53.60 ± 4.13	95.50 ± 10.89	33.87 ± 6.62	101.45 ± 3.54	
N53-2	Mean	39.00 ± 7.21	58.13 ± 5.67	54.87 ± 7.51	54.20 ± 5.64	122.60 ± 3.96	43.27 ± 6.20	184.74 ± 9.85	
DH	Mean	26.82 ± 6.98	43.02 ± 8.74	39.65 ± 10.85	54.57 ± 8.36	101.31 ± 24.09	33.07 ± 7.67	113.43 ± 27.64	
Range	Min–Max	4.3–52.1	3.6–66.3	3.5–65.1	14.3–74.9	47.7–180.8	2.9–51.4	49.0–187.8	
FBN		**08DL**	**09DL**	**10DL**	**11DL**	**11GS**	**12DL**	**12GS**	
KenC-8	Mean	13.4 ± 1.6	17.9 ± 1.7	14.7 ± 1.7	16.9 ± 0.6	8.3 ± 1.5	13.6 ± 1.4	10.5 ± 1.2	
N53-2	Mean	11.9 ± 0.9	10.4 ± 1.6	9.5 ± 0.8	8.7 ± 0.8	6.3 ± 0.7	9.4 ± 2.3	6.4 ± 1.1	
DH	Mean	14.3 ± 4.42	11.1 ± 1.73	11.3 ± 2.30	115.0 ± 2.23	7.8 ± 1.57	10.1 ± 1.58	7.2 ± 1.72	
Range	Min–Max	6.3–25.5	5.4–19.2	5.5–22.9	5.2–20.6	3.4–13.4	4.5–15.3	2.8–14.5	
LMI(cm)		**09DL**	**10DL**	**11DL**	**12DL**				
KenC-8	Mean	46.13 ± 2.62	51.94 ± 2.32	42.62 ± 5.63	50.92 ± 3.32				
N53-2	Mean	60.92 ± 5.38	65.97 ± 1.72	67.94 ± 3.45	69.54 ± 3.78				
DH	Mean	58.09 ± 6.99	55.75 ± 6.87	53.64 ± 8.02	55.27 ± 9.69				
Range	Min-Max	35.0–79.5	34.1–78.1	32.1–74.7	33.8–79.7				
PMI		**09DL**	**10DL**	**11DL**	**12DL**				
KenC-8	Mean	55.2 ± 1.8	57.2 ± 2.9	59.0 ± 3.7	49.7 ± 7.2				
N53-2	Mean	62.5 ± 4.8	66.6 ± 1.2	73.7 ± 3.0	61.7 ± 5.4				
DH	Mean	59.7 ± 8.1	57.4 ± 7.7	63.2 ± 9.1	55.4 ± 8.5				
Range	Min–Max	36.5–88.3	23.7–80.5	23.9–93.2	30.4–77.6				

[a] *Seed yield and related traits in different microenvironments.*

[b] *Mean value ± SD.*

[c] *Micro-environments.*

"KenC-8" represents male parent, "N53-2" represents female parent.

SY, seed yield; BY, biomass yield; SW, thousand seed weight; PH, plant height; BH, first effective branch height; FBN, first effective branch number; LMI, length of main inflorescence; PMI, pod number of main inflorescence.

insertion segment were considered as candidate genes for the QTL. Thirdly, the genes of *Arabidopsis* were used to identify homologous genes in *B. rapa*, *B. oleracea,* and *B. napus*. The detailed methods are found in Long et al. (2007) and Shi et al. (2009).

Ethical Standards

The authors declare that the experiments comply with the current laws of the country in which they were performed.

RESULTS

Phenotypic Analysis and Genetic Correlation Between SY and SYRTs

The SY and SYRTs of the two parents and the KN population showed differences in most micro-environments (**Table 2**). There was a wide segregation range of SY, with a continuous normal distribution and transgressive segregation in all trials (**Figure 1**), suggesting that SY was a quantitative trait with polygenic control. Seven other SYRTs (BY, SW, PH, BH, FBN, LMI, and PMI) also showed a wide segregation range in all trials with normal or near-normal distributions.

The correlations between SY and SYRTs showed large differences (**Table 3**). The results indicated that SY was highly and positively correlated with SYRTs except for FBN, and especially for BY with a correlation coefficient of 0.83. LMI was significantly positively correlated with SY and SYRTs except for FBN. SW was significantly positively correlated with SY (0.31), BY (0.53), PH (0.42), and LMI (0.37). The high correlations among SY and SYRTs indicated that these traits might be controlled by the same kinds of genes in some cases.

TABLE 3 | Pearson correlation coefficients for trait pairs affecting SY and SYRTs in KN population.

Trait	SY	BY	SW	PH	BH	FBN	LMI	PMI
SY	1	0.83**	0.31**	0.60**	0.37**	0.07	0.41**	0.49**
BY	0.83**	1	0.53**	0.74**	0.39**	0.22**	0.53**	0.37**
SW	0.31**	0.53**	1	0.42**	0.11	0.06	0.37**	0.11
PH	0.60**	0.74**	0.42**	1	0.69**	0.12*	0.74**	0.39**
BH	0.37**	0.39**	0.11	0.69**	1	0.03	0.32**	0.33**
FBN	0.07	0.22**	0.06	0.12*	0.03	1	−0.22**	−0.02
LMI	0.41**	0.53**	0.37**	0.74**	0.32**	−0.22**	1	0.35**
PMI	0.49**	0.37**	0.11	0.39**	0.33**	−0.02	0.35**	1

*$p < 0.05$, **$p < 0.01$, respectively.

Genome-Wide QTL Detection for SY and SYRTs

In total, 226 identified QTLs were observed for SY and SYRTs: 18 for SY and 208 for SYRTs (Table S2). These 226 QTLs were integrated into 144 consensus QTLs, which were located on 18 linkage groups with the exception of A8 (**Figure 2**, Figure S1).

For SY, there were nine consensus QTLs obtained, mainly located on A2, A6, C1, and C6 (**Figure 2**, Table S2). Four of these QTLs were repeatedly detected in different experiments, including *cqSY-C1-2* and *cqSY-C6-2*, detected in four successive years in DL (09DL, 10DL, 11DL, and 12DL), and *cqSY-C6-3* detected in three experiments (09DL, 10DL, and 11DL). In addition, both *cqSY-C6-2* and *cqSY-C6-3* were assumed to be major QTLs with PVE > 10% in two environments. Meanwhile, *cqSY-A2-2*, which was only observed in 10GS, was also a major QTL with PVE = 20.91% (**Table 4**).

For SYRTs, the number of QTLs for different traits clearly differed. For BY, 13 consensus QTLs were obtained, mainly located on A7, C1, and C6 with 4.34–19.96% of PVE (**Figure 2**, Table S2). Five of these QTLs were repeatedly detected in different experiments, for example, *cqBY-C1-2*, *cqBY-C6-1,* and *cqBY-C6* were detected in three experiments.

For SW, 25 consensus QTLs were detected, distributed on 11 linkage groups (**Figure 2**, Table S2). Among them, nine consensus QTLs were repeatedly detected in different environments (Table S3). QTLs *cqSW-A7-2*, *cqSW-C1-1*, and *cqSW-C1-2* were repeatedly detected in four experiments in winter area—with *cqSW-A7-2* and *cqSW-C1-1* regarded as two major QTLs with PVE > 10% in two environments (**Table 4**). The QTL *cqSW-C9-1* that appeared in both winter and spring areas was an insensitive QTL in terms of response to environments.

For PH, 18 consensus QTLs were obtained at the mature stage in eight environments (Wang et al., 2015), and mainly located on A3, C6, and C9 (**Figure 2**, Table S2). Seven QTLs were repeatedly detected in different environments (Table S3), including *cqPH-A3-3* detected in six environments, and *cqPH-C6-2* and *cqPH-C9-5* detected in three environments (Wang et al., 2015).

For BH, 27 consensus QTLs were integrated from 41 identified QTLs, and mainly located on A2, A3, A10, and C9 (Table S2). Ten QTLs were repeatedly detected in different environments (Table S3), for example, *cqBH-A3-2* and *cqBH-A10* were repeatedly observed both in winter and spring areas. Additionally, QTL *cqBH-A2* was regarded as a major QTL with PVE > 10% in 11GS and 12GS. Because BH was highly positively correlated with PH, the major QTL *cqBH-A2* for BH might also regulate PH.

For FBN, 25 consensus QTLs were obtained and were located on 12 chromosomes (Table S2). Three QTLs were repeatedly detected in both winter and spring areas: *cqFBN-A3-1*, *cqFBN-A3-2,* and *cqFBN-C3-2*. In addition, *cqFBN-A2* and *cqFBN-C2* were repeatedly detected in two and three experiments, respectively. One important QTL, *cqFBN-C6-1* with PVE > 10% in two environments, was considered as a major QTL (**Table 4**).

For LMI, 24 identified QTLs were detected and integrated into 17 consensus QTLs, of which four consensus QTLs were repeatedly detected in different experiments (Table S2). For example, *cqLMI-A3-1* and *cqLMI-A3-2* were repeatedly detected in three and four experiments, respectively. However, no major QTL was observed for LMI. For PMI, 10 consensus QTLs were obtained and only two were repeatedly detected in different environments: *cqPMI-A10* and *cqPMI-C6* in spring (11GS and 12GS) and winter areas (09DL and 11DL), respectively. No QTLs for PMI reached the standard of a major QTL.

In conclusion, with 22, linkage group A3 had the largest number of consensus QTLs, followed by C6 and C9, both with

17. More than half of the consensus QTLs (95 of 144) for SY and SYRTs were detected in one micro-environment (Table S3). Otherwise, 28, 13, and 7 consensus QTLs were identified in two, three and four micro-environments, respectively (**Figure 3**). In total, 102 and 34 consensus QTLs were detected in winter (DL) and spring areas (GS), respectively, and only eight appeared in both areas (**Figure 3**, Table S3). These results indicated that the majority of consensus QTLs were expressed principally in response to a specific environment.

The Unique QTL Analysis for SY and SYRTs

Of the 144 consensus QTLs for SY and SYRTs, 112 QTLs with overlapping CIs were integrated into 40 unique QTLs, and the remaining 32 consensus QTLs were only detected for one trait (Table S4). Altogether, 72 unique QTLs were obtained, in which 39 unique QTLs, respectively, affected 2–6 different traits. These unique QTLs were considered as pleiotropic (Table S4), such as *uq.A2-3* which was integrated from two QTLs for SY and five for SYRTs. Notably, *uq.A2-3*,

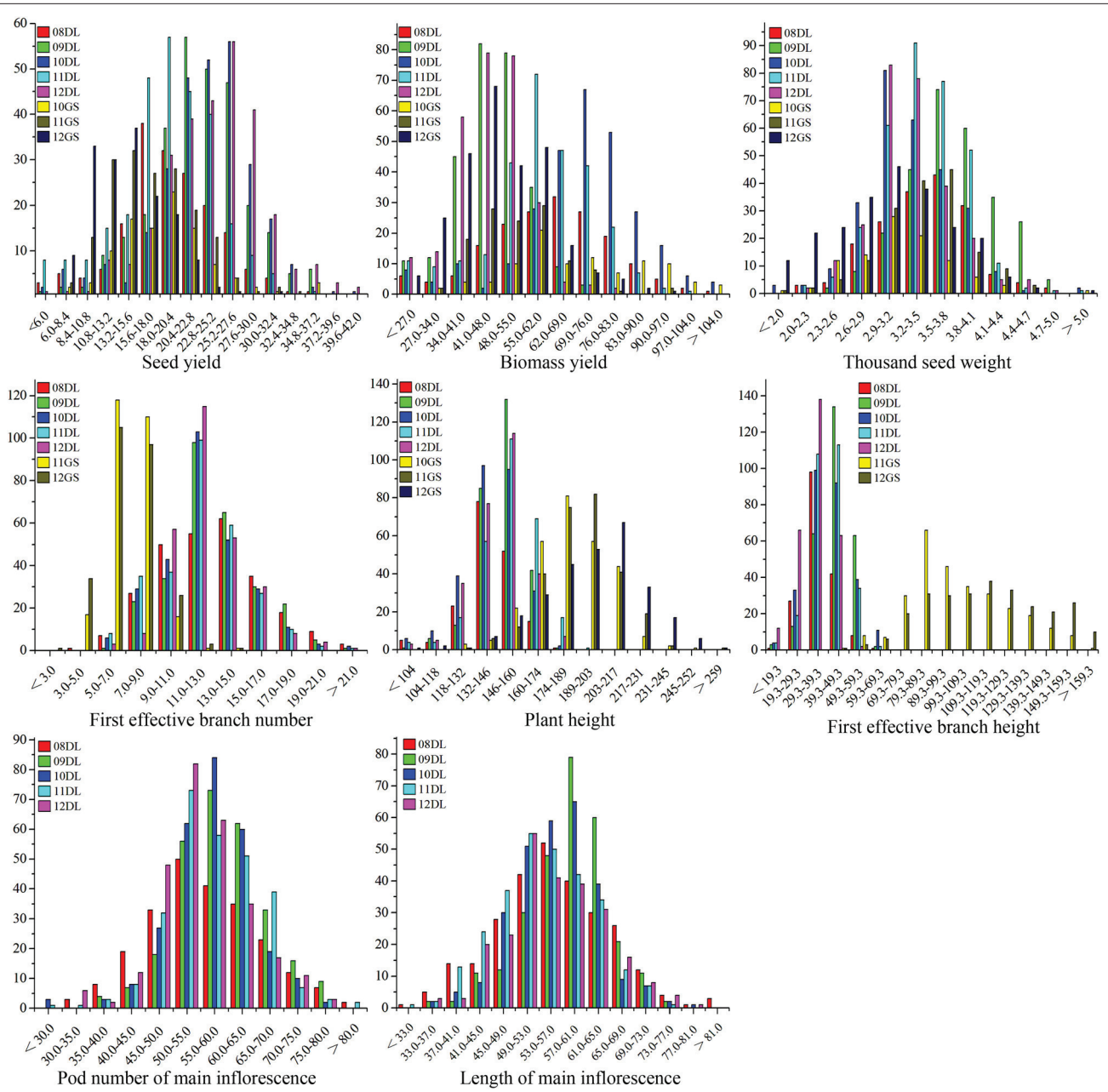


FIGURE 1 | The frequency distribution of SY and SYRTs in multiple environments. The units of the x-axis are the phenotypic values, and the units of the y-axis are the number of lines. SY and SYRTs in different experiments was discriminated using different colored boxes. A unit of measurement: seed yield (g), biomass yield (g), thousand seed weight (g), plant height (cm), first effective branch height (cm), first effective branch number, length of main inflorescence (cm), pod number of main inflorescence.

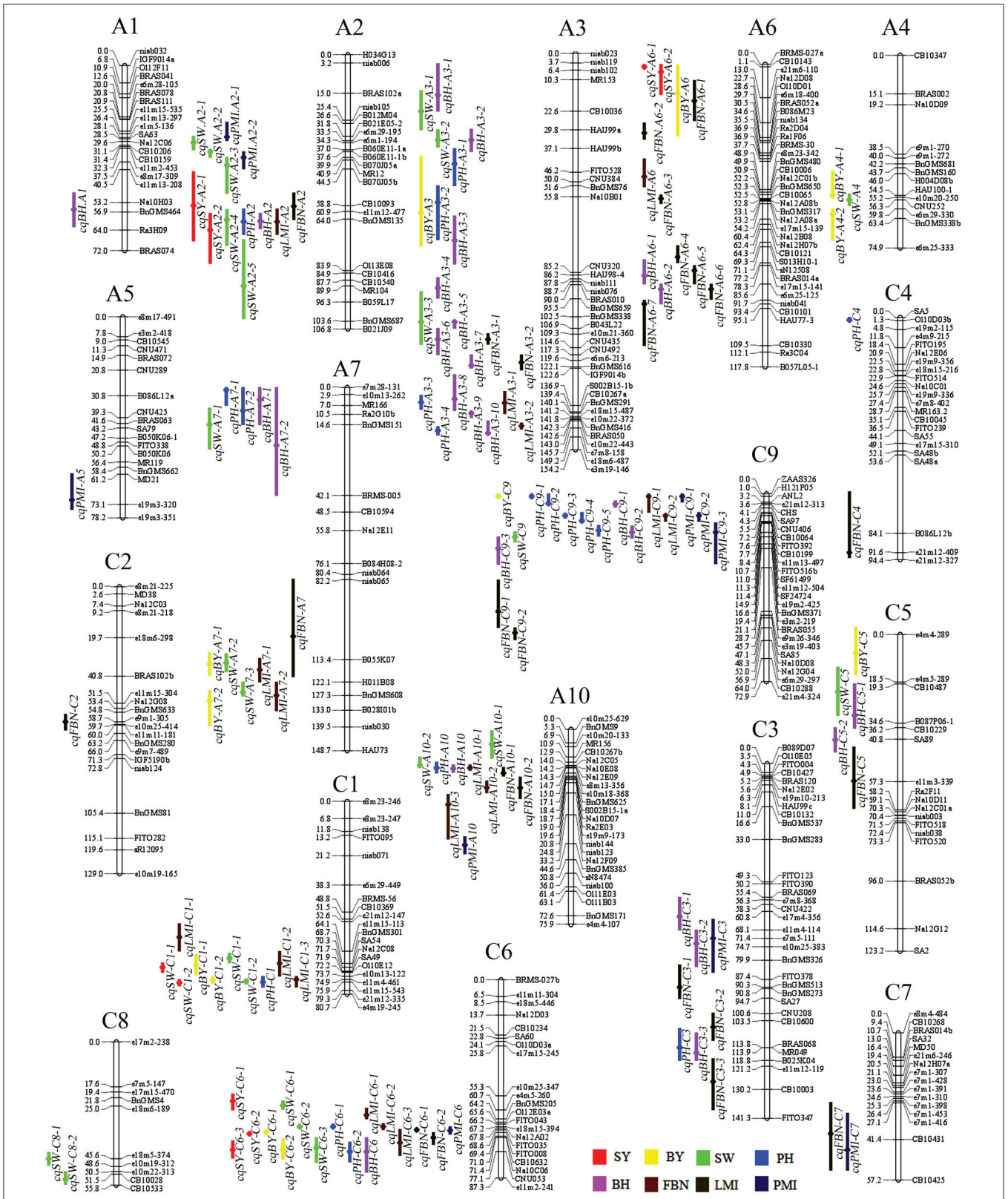


FIGURE 2 | Genetic linkage map and the location of QTLs for SY and SYRTs in the KN linkage map. The 144 consensus QTLs for SY and SYRTs were distributed on 18 linkage groups with the exception of A8, A1–A10 were represented by the A genome and C1–C9 were represented by the C genome in *B. napus*. The loci names were listed on the right of the linkage groups, while position of loci were showed on the left side of linkage groups. The consensus QTLs associated

(Continued)

FIGURE 2 | Continued
with SY and SYRTs were indicated by bars with various backgrounds on the left of each linkage group (Red bar, seed yield; Yellow bar, biomass yield; Green bar, thousand seed weight; Blue bar, plant height; Purple bar, first effective branch height; Brown bar, first effective branch number; Black bar, length of main inflorescence; Dark blue bar, pod number of main inflorescence).

TABLE 4 | Detailed information of seven major QTLs for SY and SYRTs in the KN population.

Consensus QTLs	Peak (cM)	CI[a](cM)	Identified QTLs	Chr.[b]	LOD	A[c]	PVE[d]	E[e]
cqSY-A2-2	68.11	61.0–81.2	*qSY-10GS2*	A2	5.55	−17.04	20.91	10GS
			qSY-09DL16-1	C6			5.63	09DL
cqSY-C6-2	70.06	69.33–70.79	*qSY-10DL16-1*	C6	7.40	9.53	11.69	10DL
			qSY-12DL16-2	C6			9.04	12DL
			qSY-11DL16-1	C6			13.13	11DL
cqSY-C6-3	76.14	72.85–79.42	*qSY-09DL16-2*	C6	8.78	10.06	6.76	09DL
			qSY-10DL16-2	C6			14	10DL
			qSY-11DL16-2	C6			15.92	11DL
cqSW-A7-2	114.65	111.12–118.18	*qSW-11DL7*	A7	6.9	0.16	9.99	11DL
			qSW-09DL7-2	A7			6.16	09DL
			qSW-10DL7	A7			11.92	10DL
			qSW-12DL7-1	A7			13.95	12DL
cqSW-C1-1	66.68	64.83–68.54	*qSW-08DL11-1*	C1	6.6	−0.17	19.62	08DL
			qSW-11DL11-1	C1			4.58	11DL
			qSW-08DL11-2	C1			17.83	08DL
			qSW-09DL11-1	C1			5.52	09DL
			qSW-10DL11-1	C1			5.73	10DL
cqBH-A2	64.69	61.71–67.68	*qBH-11GS2*	A2	12.7	10.45	14.58	11GS
			qBH-12GS2	A2			17.17	12GS
cqFBN-C6-1	68.6	67.86–69.35	*qFBN-09DL16*	C6	13.04	−0.91	19.14	09DL
			qFBN-10DL16	C6			13.47	10DL

[a]*Confidence interval.*
[b]*Chromosome.*
[c]*Additive.*
[d]*Phenotypic variation explained by each identified QTL.*
[e]*The environment in which QTL were detected.*
DL, Dali; GS, Sunan; 08, 09, 10, 11, and 12 indicated the years of 2008, 2009, 2010, 2011, and 2012, respectively. The QTL abbreviation "q" represents identified QTL, the QTL abbreviation "cq" represents consensus QTL.

uq.C6-5, and *uq.C6-6* were integrated from 3 to 6 QTLs of different traits and contained three major QTLs for SY (**Table 5**).

There were 21 unique QTLs observed, which were integrated from two consensus QTLs that controlled different traits (Table S4). For example, *uq.A3-1, uq.A3-2, uq.A3-6,* and *uq.C5-2* all included QTLs for SW and BH. Likewise, *uq.A2-1* and *uq.A2-2* were composed of QTLs for SW and PMI. Although *uq.A10-3* was a pleiotropic QTL and was integrated from three consensus QTLs, it was only closely linked to FBN and LMI. In addition, *uq.A4-1* controlled BY and *uq.C6-1* controlled SY, which were both closely linked to the QTL for SW. These results also explained why the QTL for SW was closely linked to the QTL for SY or BY.

Some unique QTLs were integrated from QTLs for more than two SYRTs (Table S5). QTL *uq.A3-4* controlled BY, *uq.A3-9,* and *uq.A3-11* controlled LMI, and *uq.A7-2* controlled SW, which were all, respectively, tightly linked to the BH and PH. *Uq.A7-3* controlled FBN and *uq.C1-2* controlled SY, which were both respectively closely linked to BY, SW, and LMI. It is noteworthy that *uq.A10-2* and *uq.C1-3* were closely related to five SYRTs; *uq.A10-2* controlled BH and FBN, and *uq.C1-3* controlled SY and BY, which were both tightly linked to SW, PH, and LMI.

Consensus Map Construction and QTL Comparison for SY and SYRTs among Different Mapping Populations

In the present study, five published populations (QN, SE, ER, TN, and BE) for QTL analysis of SY and SYRTs were used for consensus map construction and QTL comparison (**Table 1**, Table S6). QTLs collected in each population were first integrated into consensus QTLs using BioMercator 2.1 software. A high-density consensus map with 907 molecular markers was constructed (Figure S3). A total of 480 consensus QTLs for SY and SYRTs from five populations were obtained and 166 consensus QTLs were

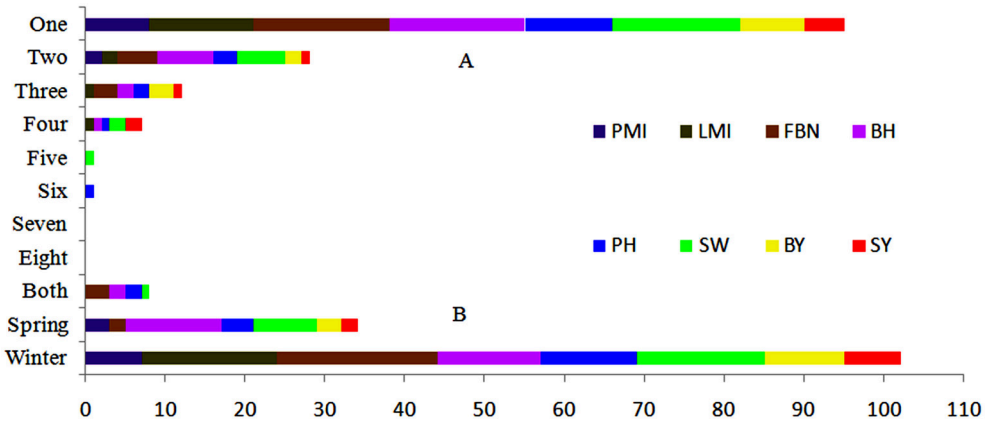


FIGURE 3 | Expression response of 144 consensus QTLs in nature environments. (A) Number of consensus QTLs appeared in one to eight micro-environments. **(B)** Number of consensus QTLs appeared in winter, spring or both macro-environments.

TABLE 5 | Detailed information of unique QTLs for SY in KN population.

Unique QTLs	Position (cM)	CI (cM)	Consensus QTLs	Chr.	Traits	A
uq.A2-3	63.95	62.0–65.9	*cqSY-A2-1*	A2	SY	−8.987
			cqSW-A2-4	A2	SW	−0.262
			cqFBN-A2	A2	FBN	−0.416
			cqLMI-A2	A2	LMI	1.975
			cqPH-A2	A2	PH	7.587
			cqBH-A2	A2	BH	10.449
			cqSY-A2-2	A2	SY	−17.037
uq.A6-1	0.03	0.0–0.6	*cqSY-A6-1*	A6	SY	8.928
			cqSY-A6-2	A6	SY	6.441
uq.C1-2	68.58	67.4–69.8	*cqLMI-C1-2*	C1	LMI	−1.742
			cqSW-C1-1	C1	SW	−0.173
			cqBY-C1-1	C1	BY	−25.009
			cqSY-C1-1	C1	SY	−7.728
uq.C1-3	75.67	75.1–76.3	*cqPH-C1*	C1	PH	−2.997
			cqLMI-C1-3	C1	LMI	−1.561
			cqBY-C1-2	C1	BY	−17.639
			cqSW-C1-2	C1	SW	−0.152
			cqSY-C1-2	C1	SY	−7.242
uq.C6-1	58.5	57.0–60.0	*cqSY-C6-1*	C6	SY	7.459
			cqSW-C6-1	C6	SW	0.150
uq.C6-5	70.23	69.5–70.9	*cqLMI-C6-3*	C6	LMI	1.899
			cqFBN-C6-2	C6	FBN	−0.637
			cqSY-C6-2	C6	SY	9.530
uq.C6-6	75.93	73.7–78.2	*cqSW-C6-3*	C6	SW	0.119
			cqBY-C6-2	C6	BY	21.017
			cqSY-C6-3	C6	SY	10.062

The QTL abbreviation "cq" represents consensus QTL, the QTL abbreviation "uq" represents unique QTL.

successfully projected onto the consensus map (**Figure 4**, Figures S2, S3).

A total of 34 QTLs for SY were projected onto the KN consensus map, half of which were located on A2, A5, A6, C2, and C6 (**Figure 5**, Figure S3, Table S7). This revealed that *KNcqSY-A2-2*, *KNcqSY-C6-1*, and five QTLs of the TN population (*TNqSY-A2-2*, *TNqSY-A2-3*, *TNqSY-A2-4*, *TNqSY-C6-1*, and *TNqSY-C6-3*) were co-localized on A2 and C6, respectively. The results indicated that these QTLs might have some important genes for SY and could be expressed stably

in different genetic backgrounds. In addition, five orthologus genes (*GASA4*, *ATCLH1*, *RBCS1A*, *LQY1,* and *ATGGH1*) for SY were aligned in the CIs of these QTLs (**Table 6**). The two QTLs (*KNcqSY-C1-1* and *KNcqSY-C1-2*) on C1 for SY were not observed in other genetic linkage groups, and these may be two new QTLs in the KN population.

For SW, 40 QTLs were projected onto the KN consensus map (**Figure 5**, Figure S3, Table S7). *KNcqSW-A2-4* and *TNqSW-A2-3* were co-localized on A2. *KNcqSW-A3-3* and five QTLs of different populations were co-localized on A3, including two QTLs of TN (*TNqSW-A3-8* and *TNqSW-A3-9*) and three of BE (*BEcqSW-A3-2*, *BEcqSW-A3-3,* and *BEcqSW-A3-4*). The co-localized QTLs on KN and TN (*KNcqSW-C6-2*, *TNqSW-C6-3*, *KNcqSW-C6-3,* and *TNqSW-C6-2*) were also observed on the C6 linkage group (Table S8). A total of five orthologous genes for SW were observed: *PTH2*, *AP2*, *LCR64*, *LCR65,* and *PDF1* (**Table 6**). These results indicated that the QTLs for SW were reliable and reproducible.

For PH, 36 QTLs were projected onto the KN consensus map, including eight and nine onto A2 and A3, respectively (**Figure 5**, Figure S3, Table S7). *KNcqPH-A2* and two QTLs of TN (*TNqPH-A2-1* and *TNqPH-A2-2*) were co-localized on A2. *KNcqPH-A3-2* and three QTLs (*TNqPH-A3-1*, *BEqPH-A3-4,* and *BEqPH-A3-3*) were co-localized on A3. Likewise, *KNcqPH-A3-3* and *TNqPH-A3-4*, *KNcqPH-C3* and *QNqHPH-C3* were also co-localized on A3 and C3, respectively. In addition, *KNcqPH-C6-1* and *TNqPH-C6-3*, *KNcqPH-C6-2* and *TNqPH-C6-2* were co-localized in the same CI of C6, respectively. The QTLs on A7 (*KNcqPH-A7-1*, *KNcqPH-A7-2,* and *KNcqSW-A7-3*), A10 (*KNcqPH-A10*), C1 (*KNcqPH-C1*), C4 (*KNcqPH-C4*), and C9 (*KNcqPH-C9-1*, *KNcqPH-C9-2*, *KNcqPH-C9-3*, *KNcqPH-C9-4,* and *KNcqPH-C9-5*) were not observed in other genetic linkage groups. One orthologous gene *TGH* (*Bra035958*) on A6 was observed (**Table 6**), but no gene was aligned in the CI of QTLs for PH.

There were 34 QTLs for FBN projected onto the KN consensus map (**Figure 5**, Figure S3, Table S7), of which 10 were projected onto A3, and only two (*KNcqFBN-A3-2* and *BEqBN-A3-4*) were co-located (Table S8). Two QTLs were projected onto A6, of which *KNcqFBN-A6-3* and *TNqBN-A6*, and *KNcqFBN-A6-5* and *SEfb6.1* were co-localized. Three QTLs of KN (*KNcqFBN-C3-1*, *KNcqFBN-C3-2*, *KNcqFBN-C3-3*) and four QTLs of QN (*QNqHFB-C3-2*, *QNqHFB-C3-1*, *QNqFB-C3-1,* and *QNqFB-C3-2*) were co-localized. In addition, *KNcqFBN-C5* and *QNqFB-C5*, and *KNcqFBN-C6-1* and *TNqBN-C6-1* were co-localized on C5 and C6, respectively. Some potential new QTLs of the KN population were obtained on A2 (*KNcqFBN-A2*), A7 (*KNcqFBN-A7-1*), A10 (*KNcqFBN-A10-1* and *KNcqFBN-A10-2*), C2 (*KNcqFBN-C2*), C4 (*KNcqFBN-C4*), C7 (*KNcqFBN-C7*), and C9 (*KNcqFBN-C9-1* and *KNcqFBN-C9-2*).

The traits BY, BH, LMI, and PMI have rarely been studied in other populations, thus, no QTL for PMI was projected onto the KN consensus map. Due to lack of common markers between KN and other maps, only a few QTLs for BY, BH, LMI, and PMI from five populations were projected onto the KN consensus map, including six QTLs for BH and eight each for BY and LMI (Figure S3, Table S8), respectively. However, none of these QTLs were co-located with the QTLs in KN population.

TABLE 6 | Positions of homologous genes of *Arabidopsis* mapped on the KN linkage map and orthologous genes between *B. napus* and *B. rapa/B. oleracea*.

Traits	Gene model	*Arabidopsis* genes	Chr.	Gene position (cM)	QTLs	QTL position (cM)	CI. (cM)	Genes of *B. rapa/B. oleracea*	Genes of *B. napus*
SY	*AT5G15230*	*GASA4*	A2	26.56				*Bra023513*	*BnaA02g02560D*
	AT1G19670	*ATCLH1*	A6	53.14				*Bra025756*	BnaA06g13830D
	AT1G67090	*RBCS1A*	C6	68.58					
	AT1G75690	*LQY1*	C6	71.37	*KNcqSY-C6-2*	70.06	65.64–77.08	*Bol027707*	BnaC06g36350D
	AT1G78660	*ATGGH1*	C6	77.08	*KNcqSY-C6-3*	76.14	65.64–77.08	*Bol027466*	BnaC06g39090D
SW	*AT5G10700*	*PTH2*	A2	3.16	*KNcqSW-A2-1*	34.31	0.0–44.0	*Bra006072*	
					KNcqSW-A2-2	37.61			
					KNcqSW-A2-3	43.01			
	AT4G36920	*AP2*	A3	98.68	*KNcqSW-A3-3*	104.15	92.8–113.3	*Bra011741*	*BnaA03g53830D*
	AT1G73603	*LCR64*	C6	71.37					
	AT1G73607	*LCR65*	C6	71.37	*KNcqSW-C6-2,*	67.2	65.16–78.92		
	AT1G75830	*PDF1*	C6	71.37	*KNcqSW-C6-3*	75.41	65.16–78.92	*Bol039878*	BnaC06g22110D
BY	*AT2G36450*	*HARDY*	A4	46.02	*KNcqBY-A4-1*	54.51	45.0–54.8	*Bra017235*	
PH	*AT5G23080*	*TGH*	A6	62.36			60.88–63.84	*Bra035958*	BnaC09g05000D
BH	*AT3G52430*	*ATPAD4*	A2	96.28			94.8–97.76	*Bra006922*	

The QTL abbreviation "KNcq" represents consensus QTL from the KN Population.

In addition, two orthologous genes were observed, including *HARDY* (*Bra017235*) for BY and *ATPAD4* (*Bra006922*) for BH (**Table 6**).

DISCUSSION

SY and SYRTs for *B. napus* are complex quantitative traits and easily affected by the environment (Quarrie et al., 2006). In previous studies, Chen et al. (2007) obtained 52 QTLs for six SYRTs and Fan et al. (2010) identified nine QTLs for SW. Chen et al. (2010) obtained 18 QTLs for SY and 22 QTLs for flowering time, while Butruille et al. (1999) found two QTLs for SY. These previous studies did not examine the relationship between SY and SYRTs, this was considered in the present study, showing correlation coefficients in the range of 0.31–0.83, except for FBN.

Several studies have revealed that only a few QTLs for SYRTs were stable in different environments (Li et al., 2007; Shi et al., 2009). The majority of consensus QTLs were also environment specific in the present study, including 70.8 and 23.6% of QTLs there were only detected in winter and spring areas, respectively. Only a few QTLs (5.6%) were detected in both areas, consistent with previous reports. These findings will be helpful for breeders to develop varieties with special adaptability.

Trait-by-trait meta-analysis revealed that 144 consensus QTLs for SY and SYRTs were integrated into 72 pleiotropic unique QTLs. Among them, six of seven unique QTLs for SY were co-located with two to five QTLs for SYRTs. On average, one unique QTL for SY involved 2.5 QTLs for SYRTs (**Table 5**). For example, *uq.A2-3* affected SY as well as five SYRTs: SW, BH, PH, FBN, and LMI. These findings were consistent with the strong correlations

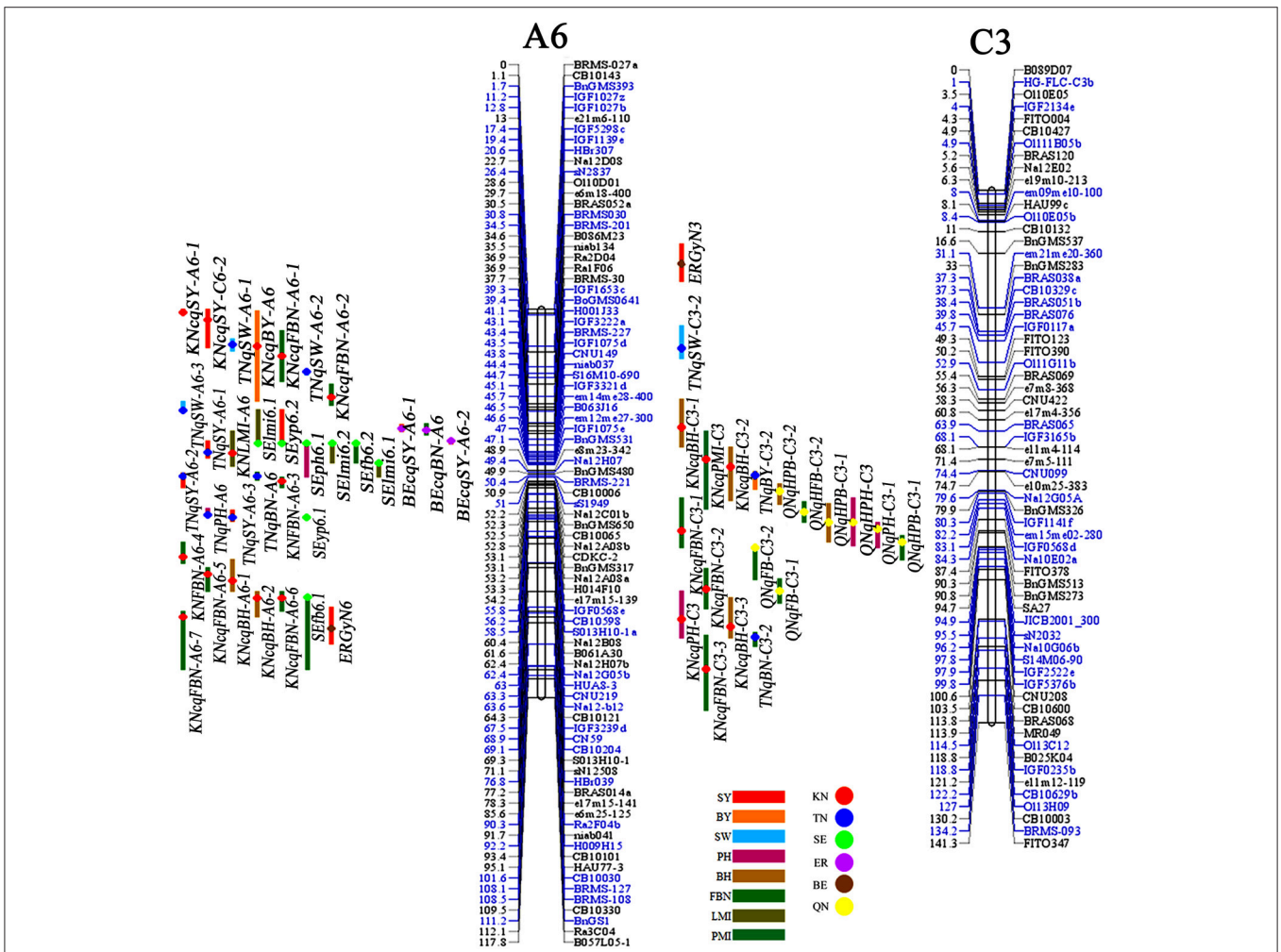


FIGURE 4 | The consensus map and QTLs for SY and SYRTs in different populations (A6 and C3). Markers with blue color indicated these makers were projected from other maps on the KN map based on common markers by BioMercator 2.1 software. QTLs for SY and SYRTs detected in different populations were discriminated with different color bars on the left of each linkage group. Red bar, SY (seed yield); Orange bar, BY (biomass yield); Cambridge blue bar, SW (thousand seed weight); Purple bar, PH (plant height); Claybank bar, BH (first effective branch height); Green bar, FBN (first effective branch number); Breen bar, LMI (length of main inflorescence); Blue-green bar, PMI (pod number of main inflorescence). The KN population and five populations were indicated by disks with various backgrounds on the bars of each QTL. Red disk, KN population; Blue disk, TN population; Light green disk, SE population; Light purple disk, ER population; Brown disk, BE population; Yellow disk, QN population.

between SY and SYRTs. Shi et al. (2009) demonstrated that the QTLs for SY were pleiotropic and synthesized, and numerous SYRTs were potential contributor to tightly link with the QTLs for SY. Li et al. (2007) revealed that QTLs for SY and SYRTs usually had overlapping regions. Similar results were also reported for spring barley (Li et al., 2005) and red clover (Herrmann et al., 2006). The genes identified as having a function for SY and SYRTs also had pleiotropic effects for at least one trait (Ashikari et al., 2005; Lim et al., 2014) or multiple traits (Li et al., 1997; Hall et al., 2005; Quarrie et al., 2006; Burgess-Herbert et al., 2008; Hu et al., 2008). In other words, QTLs for SY and SYRTs might have resulted from pleiotropic QTLs that controlled multiple traits by containing multiple, closely linked, trait-specific genes (Hall et al., 2005).

Indicator QTLs have been successfully used in identifying genes with pleiotropic effects for SY (Shi et al., 2009). Indicator QTLs for SY must be stably expressed, easily measured, and identifiable candidate genes than the co-localized QTLs for SY. Using the similar methods, the candidate genes were successfully cloned by the QTL for BY in wheat (Quarrie et al., 2006), and the QTL for flowering time could be regarded as an indicator QTL for natural variation in rice (Xue et al., 2008). In the present study, four QTLs for SY closely co-existed with QTLs for SW (**Table 5**). SW can be more easily and precisely measured than SY, and is also less influenced by environmental factors than other SYRTs (Shi et al., 2009; Ding et al., 2012). Therefore, QTLs for SW could be regarded as indicator QTLs for SY. In future research, cloning the indicator QTL *KNcqSW-C6-3* for SY is more feasible than cloning *KNcqSY-C6-3*.

In the present study, 166 QTLs for SY and SYRTs from five populations were projected onto the constructed high-density consensus map. However, only 36 QTLs were co-located with QTLs identified in the KN population, including six QTLs for SY from the TN population and 11 for SW from the TN and BE population. Meanwhile, the co-located QTLs for SY were located on A2 and C6, while the co-located QTLs for SW were mainly located on A3 and C6. Thus, C6 was an important linkage group and included important genes for SY and SW. The reason that only a few QTLs were co-localized might be the lack of sufficient common markers between the different maps (Zhou et al., 2014) and also that there has been little research for some traits, such as PMI and LMI.

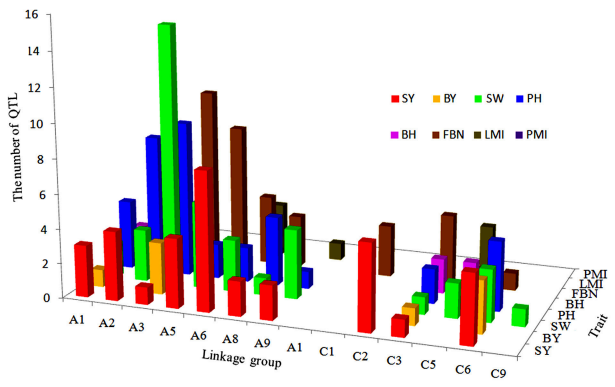


FIGURE 5 | The number and distribution of projection QTLs for SY and SYRTs from five population in *B. napus*. The x-axis of abscissas indicated 19 linkage groups, y-axis indicated the number of projection QTLs and z-axis indicated SY and SYRTs.

The known *Arabidopsis* genome sequence has been exploited as a tool for comparative analysis between *Arabidopsis* and *Brassicaceae* genomes, and the conserved genomic block has been identified in different *Brassicaceae* species (Arcade et al., 2004; Boivin et al., 2004; Schranz et al., 2006). This provided a method to align candidate genes underlying QTLs controlling important agronomic traits (Long et al., 2007). For SY, the orthologous gene *LQY1* was mapped on the C6 chromosome, which encodes a small zinc-finger-containing thylakoid membrane protein of *Arabidopsis* (Lu, 2011; Jin et al., 2014). Gene *LQY1* resulted in a lower quantum yield of photosystem II (PSII) photochemistry and reduced PSII electron transport rate following high-light treatment (Lu et al., 2011). For SW, gene *AP2* (*Bra011741* and *BnaA03g53830D*) and *PDF1* (*Bol039878* and *BnaC06g22110D*) were underlying the CIs of *KNcqSW-A3-3* and *KNcqSW-C6-3*, respectively. Gene *AP2* was involved in the specification of floral organ identity, establishment of floral meristem identity, ovule and seed coat development, and also had a role in controlling seed mass (Ohto et al., 2009). Gene *PDF1* is known as plant defense type 1 gene, and conferred high capacities to tolerate and hyperaccumulate zinc and cadmium (Mirouze et al., 2006). In this study, gene *PDF1* located on C6 controlled seed size and weight; however, the genetic mechanisms of controlling seed size and weight remained ambiguous. These genes are speculated to be candidate genes for SY or SYRTs in *B. napus*.

CONCLUSION

The genetic mechanisms underlying SY and SYRTs were analyzed through QTL analysis in *B. napus*. A total of 226 QTLs were identified, and were integrated into 144 consensus QTLs. Seven major QTLs were obtained, including three QTLs for SY, two for SW and one each for BH and FBN, respectively. Trait-by-trait meta-analysis revealed that one unique QTL for SY involved 2.5 QTLs for SYRTs. Meanwhile, QTL projection from five different genetic linkage maps onto the KN consensus map showed that 36 QTLs were co-localized with QTLs identified in the KN population. In addition, candidate genes for SY and SYRTs were observed, including five each for SY (*GASA4*, *ATCLH1*, *RBCS1A*, *LQY1,* and *ATGGH1*) and SW (*PTH2*, *AP2*, *LCR64*, *LCR65,* and *PDF1*), and one each for BY (*HARDY*), BH (*ATPAD4*), and PH (*TGH*). The obtained candidate genes for SY and SYRTs were conducive to fine mapping and key gene cloning. These findings will be valuable in hybrid cultivar breeding and in analyzing QTL expression in different environments.

AUTHOR CONTRIBUTIONS

WZ and XW carried out the QTL analysis and wrote the manuscript. JT, BL, LC, and HC participated in the field experiment. YL, JX, JG, and WL made helpful suggestions to

the manuscript. HW and ML designed, led and coordinated the overall study.

FUNDING

This work was supported financially by the National Science Foundation of China (31471532), the National Basic Research Program of China (2015CB150205), International Cooperation in Science and Technology Projects (2014DFA32210), the New Century Talents Support Program of the Ministry of Education of China (NCET110172), the Natural Science Foundation Research Key Projects of Shaanxi Province of China (2013JZ005) and International Cooperation Key Projects of Shaanxi Province of China (2012KW-15).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00017

REFERENCES

Arcade, A., Labourdette, A., Falque, M., Mangin, B., Chardon, F., Charcosset, A., et al. (2004). BioMercator: integrating genetic maps and QTL towards discovery of candidate genes. *Bioinformatics* 20, 2324–2326. doi: 10.1093/bioinformatics/bth230

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., et al. (2005). Cytokinin oxidase regulates rice grain production. *Science* 309, 741–745. doi: 10.1126/science.1113373

Basunanda, P., Radoev, M., Ecke, W., Friedt, W., Becker, H. C., and Snowdon, R. J. (2010). Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theor. Appl. Genet.* 120, 271–281. doi: 10.1007/s00122-009-1133-z

Boivin, K., Acarkan, A., Mbulu, R. S., Clarenz, O., and Schmidt, R. (2004). The Arabidopsis genome sequence as a tool for genome analysis in Brassicaceae. A comparison of the *Arabidopsis* and capsella rubella genomes. *Plant Physiol.* 135, 735–744. doi: 10.1104/pp.104.040030

Burgess-Herbert, S., Cox, A., Tsaih, S. W., and Paigen, B. (2008). Practical applications of the bioinformatics tool box for narrowing quantitative trait loci. *Genetics* 180, 2227–2235. doi: 10.1534/genetics.108.090175

Butruille, D. V., Guries, R. P., and Osborn, T. C. (1999). Linkage analysis of molecular markers and quantitative trait loci in populations of inbred backcross lines of *Brassica napus* L. *Genetics* 153, 949–964.

Cai, D., Xiao, Y., Yang, W., Ye, W., Wang, B., Younas, M., et al. (2014). Association mapping of six yield-related traits in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 127, 85–96. doi: 10.1007/s00122-013-2203-9

Chen, G., Geng, J., Rahman, M., Liu, X., Tu, J., Fu, T., et al. (2010). Identification of QTL for oil content, seed yield, and flowering time in oilseed rape (*Brassica napus*). *Euphytica* 175, 161–174. doi: 10.1007/s10681-010-0144-9

Chen, W., Zhang, Y., Liu, X., Chen, B., Tu, J., Fu, T., et al. (2007). Detection of QTL for six yield-related traits in oilseed rape (*Brassica napus*) using DH and immortalized F_2 populations. *Theor. Appl. Genet.* 115, 849–858. doi: 10.1007/s00122-007-0613-2

Ding, G., Liao, Y., Yang, M., Zhao, Z., Shi, L., and Xu, F. (2011). Development of gene-based markers from *Arabidopsis thaliana* functional genes involved in phosphorpus homeostasis and mapping in *Brassica napus*. *Euphytica* 181, 305–322. doi: 10.1007/s10681-011-0428-8

Ding, G., Zhao, Z., Liao, Y., Hu, Y., Shi, L., Long, Y., et al. (2012). Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in *Brassica napus*. *Ann. Bot.* 109, 747–759. doi: 10.1093/aob/mcr323

Fan, C., Cai, G., Qin, J., Li, Q., Yang, M., Wu, J., et al. (2010). Mapping of quantitative trait loci and development of allele-specific markers for seed weight in *Brassica napus*. *Theor. Appl. Genet.* 121, 1289–1301. doi: 10.1007/s00122-010-1388-4

Hall, M. C., Basten, C. J., and Willis, J. H. (2005). Pleiotropic quantitative trait loci contribute to population divergence in traits associated with life-history variation in mimulus guttatus. *Genetics* 172, 1829–1844. doi: 10.1534/genetics.105.051227

Herrmann, D., Boller, B., Studer, B., Widmer, F., and Kolliker, R. (2006). QTL analysis of seed yield components in red clover (*Trifolium pratense* L.). *Theor. Appl. Genet.* 112, 536–545. doi: 10.1007/s00122-005-0158-1

Hu, K., Qiu, D., Shen, X., Li, X., and Wang, S. (2008). Isolation and manipulation of quantitative trait loci for disease resistance in rice using a candidate gene approach. *Mol. Plant* 1, 786–793. doi: 10.1093/mp/ssn039

Jiang, C., Shi, J., Li, R., Long, Y., Wang, H., Li, D., et al. (2014).Quantitative trait loci that control the oil content variation of rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 127, 957–968. doi: 10.1007/s00122-014-2271-5

Jiang, C., and Zeng, Z. (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* 140, 1111–1127.

Jin, H., Liu, B., Luo, L., Feng, D., Wang, P., Liu, J., et al. (2014). Hypersensitive to high light1 interacts with low quantum yield of photosystem II1 and functions in protection of photosystem II from photodamage in *Arabidopsis*. *Plant Cell* 26, 1213–1229. doi: 10.1105/tpc.113.122424

Kearsey, M. J., and Pooni, H. S. (1998). *The Genetical Analysis of Quantitative Traits.* Cheltenham: Stanley Thornes Publishers Ltd.

Li, J., Huang, X., Heinrichs, F., Ganal, M. W., and Roder, M. S. (2005). Analysis of QTLs for yield, yield components, and malting quality in a BC_3-DH population of spring barley. *Theor. Appl. Genet.* 110, 356–363. doi: 10.1007/s00122-004-1847-x

Li, Y., Shen, J., Wang, T., Chen, Q., Zhang, X., Fu, T., et al. (2007). QTL analysis of yield-related traits and their association with functional markers in *Brassica napus* L. *Aust. J. Agr. Res.* 58, 759. doi: 10.1071/AR06350

Li, Z., Pinson, S. R., Park, W. D., Paterson, A. H., and Stansel, J. W. (1997). Epistasis for three grain yield components in rice (*Oryza sativa* L.). *Genetics* 145, 453–465.

Lim, J., Yang, H., Jung, K., Yoo, S. C., and Paek, N. C. (2014). Quantitative trait locus mapping and candidate gene analysis for plant architecture traits using whole genome re-sequencing in rice. *Mol. Cells* 37, 149–160. doi: 10.14348/molcells.2014.2336

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5:3930. doi: 10.1038/ncomms4930

Lombard, V., and Delourme, R. (2001). A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theor. Appl. Genet.* 103, 491–507. doi: 10.1007/s001220100560

Long, Y., Shi, J., Qiu, D., Li, R., Zhang, C., Wang, J., et al. (2007). Flowering time quantitative trait loci analysis of oilseed *Brassica* in multiple environments and genomewide alignment with *Arabidopsis*. *Genetics* 177, 2433–2444. doi: 10.1534/genetics.107.080705

Lu, Y. (2011). The occurrence of a thylakoid-localized small zinc finger protein in land plants. *Plant Signal. Behav.* 6, 1881–1885. doi: 10.4161/psb.6.12.18022

Lu, Y., Hall, D. A., and Last, R. L. (2011). A small Zinc finger thylakoid protein plays a role in maintenance of Photosystem II in *Arabidopsis thaliana*. *Plant Cell* 23, 1861–1875. doi: 10.1105/tpc.111.085456

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297

Mirouze, M., Sels, J., Richard, O., Czernic, P., Loubet, S., Jacquier, A., et al. (2006). A putative novel role for plant defensins: a defensin from the zinc hyper-accumulating plant, *Arabidopsis* halleri, confers zinc tolerance. *Plant J.* 47, 329–342. doi: 10.1111/j.1365-313X.2006.02788.x

Ohto, M. A., Floyd, S. K., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2009). Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in *Arabidopsis*. *Sex. Plant Reprod.* 22, 277–289. doi: 10.1007/s00497-009-0116-1

Paran, I., and Zamir, D. (2003). Quantitative traits in plants: beyond the QTL. *Trends Genet.* 19, 303–306. doi: 10.1016/S0168-9525(03)00117-3

Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., et al. (2006). A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theor. Appl. Genet.* 114, 67–80. doi: 10.1007/s00122-006-0411-2

Quarrie, S., Pekic, Q. S., Radosevic, R., Kaminska, A., Barnes, J. D., Leverington, M., et al. (2006). Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. *J. Exp. Bot.* 57, 2627–2637. doi: 10.1093/jxb/erl026

Quijada, P. A., Udall, J. A., Lambert, B., and Osborn, T. C. (2006). Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 1. Identification of genomic regions from winter germplasm. *Theor. Appl. Genet.* 113, 549–561. doi: 10.1007/s00122-006-0323-1

Qzer, H., Oral, E., and Dogru, U. (1999). Relationships between yield and yield components on currently improved spring rapeseed cultivars. *Tr. J. Agric. For.* 53, 603–607.

Radoev, M., Becker, H. C., and Ecke, W. (2008). Genetic analysis of heterosis for yield and yield components in Rapeseed (*Brassica napus* L.) by quantitative trait locus mapping. *Genetics* 179, 1547–1558. doi: 10.1534/genetics.108.089680

Raman, H., Raman, R., Eckermann, P., Coombes, N., Manoli, S., Zou, X., et al. (2013). Genetic and physical mapping of flowering time loci in canola (*Brassica napus* L.). *Theor. Appl. Genet.* 126, 119–132. doi: 10.1007/s00122-012-1966-8

Schranz, M. E., Lysak, M. A., and Mitchell-Olds, T. (2006). The ABC's of comparative genomics in the *Brassicaceae*: building blocks of crucifer genomes. *Trends Plant Sci.* 11, 535–542. doi: 10.1016/j.tplants.2006.09.002

Scoles, G., Fu, F., Liu, L., Chai, Y., Chen, L., Yang, T., et al. (2007). Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments. *Genome* 50, 840–854. doi: 10.1139/G07-068

Shi, J., Li, R., Qiu, D., Jiang, C., Long, Y., Morgan, C., et al. (2009). Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. *Genetics* 182, 851–861. doi: 10.1534/genetics.109.101642

Shi, L., Shi, T., Broadley, M. R., White, P. J., Long, Y., Meng, J., et al. (2013). High-throughput root phenotyping screens identify genetic loci associated with root architectural traits in *Brassica napus* under contrasting phosphate availabilities. *Ann. Bot.* 112, 381–389. doi: 10.1093/aob/mcs245

Udall, J. A., Quijada, P. A., Lambert, B., and Osborn, T. C. (2006). Quantitative trait analysis of seed yield and other complex traits in hybrid spring rapeseed (*Brassica napus* L.): 2. Identification of alleles from unadapted germplasm. *Theor. Appl. Genet.* 113, 597–609. doi: 10.1007/s00122-006-0324-0

UN (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot.* 7, 389–452.

Wang, S., Basten, C., and Zeng, Z. (2007). *Windows QTL Cartographer 2.5.* Raleigh, NC: North Carolina State University.

Wang, X., Wang, H., Long, Y., Li, D., Yin, Y., Tian, J., et al. (2013). Identification of QTLs associated with oil content in a high-oil *Brassica napus* cultivar and construction of a high-density consensus map for QTLs comparison in *B. napus*. *PLoS ONE* 8:e80569. doi: 10.1371/journal.pone.0080569

Wang, X., Wang, H., Long, Y., Liu, L., Zhao, Y., Tian, J., et al. (2015). Dynamic and comparative QTL analysis for plant height in different developmental stages of *Brassica napus* L. *Theor. Appl. Genet.* 128, 1175–1192. doi: 10.1007/s00122-015-2498-9

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., et al. (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767. doi: 10.1038/ng.143

Yu, F. (1998). Genetic analyses of several quantitative traits of doubled haploid population in *Brassica napus* L. (in Chinese with an English abstract). *Sci. Agric. Sinica* 31, 44–48.

Zhang, G., He, Y., Xu, L., Tang, G., and Zhou, W. (2006). Genetic analyses of agronomic and seed quality traits of doubled haploid population in *Brassica napus* through microspore culture. *Euphytica* 149, 169–177. doi: 10.1007/s10681-005-9064-5

Zhou, Q., Fu, D., Mason, A., Zeng, Y., Zhao, C., and Huang, Y. (2014). *In silico* integration of quantitative trait loci for seed yield and yield-related traits in *Brassica napus*. *Mol. Breeding* 33, 881–894. doi: 10.1007/s11032-013-0002-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhao, Wang, Wang, Tian, Li, Chen, Chao, Long, Xiang, Gan, Liang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 July 2016
doi: 10.3389/fpls.2016.01058

Multigenic Control of Pod Shattering Resistance in Chinese Rapeseed Germplasm Revealed by Genome-Wide Association and Linkage Analyses

Jia Liu[1†], Jun Wang[1,2†], Hui Wang[1], Wenxiang Wang[1], Rijin Zhou[1], Desheng Mei[1], Hongtao Cheng[1], Juan Yang[1], Harsh Raman[3] and Qiong Hu[1*]*

[1] Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China, [2] Graduate School of Chinese Academy of Agricultural Sciences, Beijing, China, [3] Graham Centre for Agricultural Innovation (an Alliance between NSW Department of Primary Industries and Charles Sturt University), Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia

OPEN ACCESS

Edited by:
Joshua L. Heazlewood, University of Melbourne, Australia

Reviewed by:
Elisa Bellucci, Marche Polytechnic University, Italy
Marie Bruser, John Innes Centre, UK

***Correspondence:**
Harsh Raman
harsh.raman@dpi.nsw.gov.au
Qiong Hu
huqiong01@caas.cn

[†] *These authors have contributed equally to this work.*

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *29 January 2016*
Accepted: *06 July 2016*
Published: *21 July 2016*

Citation:
Liu J, Wang J, Wang H, Wang W, Zhou R, Mei D, Cheng H, Yang J, Raman H and Hu Q (2016) Multigenic Control of Pod Shattering Resistance in Chinese Rapeseed Germplasm Revealed by Genome-Wide Association and Linkage Analyses. Front. Plant Sci. 7:1058. doi: 10.3389/fpls.2016.01058

The majority of rapeseed cultivars shatter seeds upon maturity especially under hot-dry and windy conditions, reducing yield and gross margin return to growers. Here, we identified quantitative trait loci (QTL) for resistance to pod shatter in an unstructured diverse panel of 143 rapeseed accessions, and two structured populations derived from bi-parental doubled haploid (DH) and inter-mated (IF_2) crosses derived from R1 (resistant to pod shattering) and R2 (prone to pod shattering) accessions. Genome-wide association analysis identified six significant QTL for resistance to pod shatter located on chromosomes A01, A06, A07, A09, C02, and C05. Two of the QTL, *qSRI.A09* delimited with the SNP marker Bn-A09-p30171993 (A09) and *qSRI.A06* delimited with the SNP marker Bn-A06-p115948 (A06) could be repeatedly detected across environments in a diversity panel, DH and IF_2 populations, suggesting that at least two loci on chromosomes A06 and A09 were the main contributors to pod shatter resistance in Chinese germplasm. Significant SNP markers identified in this study especially those that appeared repeatedly across environments provide a cost-effective and an efficient method for introgression and pyramiding of favorable alleles for pod shatter resistance via marker-assisted selection in rapeseed improvement programs.

Keywords: rapeseed, pod shatter resistance, genetic linkage mapping, genome-wide association, design breeding

INTRODUCTION

Rapeseed (*Brassica napus* L., 2n = 4× = 38, genome AACC) is the third largest oilseed crop produced in the world after oil palm and soybean (USDA FAS, 2015)[1]. In nature, many plant species including rapeseed dehisce seeds easily upon maturity for dispersal and survival in subsequent generations. However, this phenomenon is one of the major bottlenecks in rapeseed production on a commercial scale. The yield loss due to seed shatter usually accounts for about 5–10% of total production; and under relatively harsh climatic conditions, it can reach up to 50% (Kadkol et al., 1984; Price et al., 1996). Moreover, shattered seeds become "volunteers" in subsequent crops in the

[1] Oilseeds: World Markets and Trade|USDA FAS. Fas.usda.gov. Retrieved 2015-08-25.

rotation cycle, making crop management difficult and expensive (Morgan et al., 2000). Rapeseed is generally harvested by windrowing or swathing. However, in recent years, farmers prefer to use combine harvesters, as this operation is less-labor intensive and cheaper compared to windrowing and manual harvesting. The latter is not an option for many western countries where rapeseed is often used as a broad-acre crop and harvested under very hot and dry conditions. Therefore, developing pod shatter resistant varieties suitable for combine harvesting has become one of the main breeding objectives of rapeseed improvement programs.

A limited genetic variation exists for pod shatter resistance in natural germplasm of rapeseed (Morgan et al., 1998; Wen et al., 2008). For example, Wen et al. (2008) evaluated 229 genotypes of rapeseed and identified only two genotypes having moderate levels of resistance to pod shatter. However, genetic variation for higher levels of resistance to pod shatter is present in other close relatives of rapeseed, such as *Brassica rapa*, *Brassica juncea,* and *Brassica carinata* (Kadkol et al., 1984; Mongkolporn et al., 2003; Raman et al., 2014). These related species have been utilized to improve pod shatter resistance in rapeseed via interspecific hybridization (Liu, 1994; Wei et al., 2010; Raman et al., 2014).

To gain insight into the genetic basis underlying quantitative variation in traits of agricultural significance such as pod shatter resistance and to enhance predictive selection efficiency in plant breeding programs, genetic mapping has become an important tool (Mauricio, 2001). Recent developments in next-generation sequencing technology, discovery of high throughput marker systems such as high density SNP markers (Trick et al., 2009; Bancroft et al., 2011), genotyping-by-sequencing (Raman et al., 2014; Bayer et al., 2015) and sequence capture (Schiessl et al., 2014), availability of chromosome based sequence of *B. rapa, B. oleracea, and B. napus* genomes (Wang et al., 2011; Chalhoub et al., 2014; Liu et al., 2014; Parkin et al., 2014) and bioinformatics, have enabled improving genomic selection of desirable alleles through marker-assisted selection in rapeseed. Multigenic inheritance for pod shatter resistance has been reported in *B. rapa*, and *B. napus* (Kadkol et al., 1986; Hossain et al., 2011; Wen et al., 2013). During the last 5 years, up to 10 QTL associated with resistance to pod shatter have been identified in both genetic mapping populations derived from doubled haploid (DH) lines (Hu et al., 2012; Wen et al., 2013; Raman et al., 2014) and a diversity panel of rapeseed accessions, originated mainly from Australia (Raman et al., 2014). Genetic loci associated with pod shatter resistance has also been mapped in *B. rapa* using RAPD markers (Mongkolporn et al., 2003), and soybean (Gao and Zhu, 2013). Several genes such as *IND*, *ALC*, *SHP1*, *SHP2*, and *FUL* and their complex regulatory network involved in pod dehiscence have been identified in *Arabidopsis,* rice and soybean (Ferrándiz et al., 2000; Liljegren et al., 2000; Rajani and Sundaresan, 2001; Konishi et al., 2006; Lewis et al., 2006; Li et al., 2006; Østergaard, 2009; Zhou et al., 2012; Dong et al., 2014; Funatsuki et al., 2014; Yoon et al., 2014).

In this study, we performed a genome wide association study (GWAS) in a diversity panel of 143 accessions and classical QTL analyses utilizing a DH population and inter-mated F_2 (IF_2) population derived from R1 (resistant to pod shatter) and R2 (prone to pod shatter) rapeseed advanced breeding lines of Chinese origin to identify loci involved in pod shatter resistance. The publicly available 60K Brassica Infinium® SNP array was utilized to genotype mapping populations. We uncovered that pod shatter resistance is controlled by multiple loci having both major and minor allelic effects. Identification of loci via GWAS and classical QTL analyses, and SNP marker significantly associated with pod shatter resistance may facilitate a cost-effective marker assisted selection of favorable alleles in rapeseed breeding programs.

MATERIALS AND METHODS

Association Mapping Population

A total of 143 diverse rapeseed accessions including 6 elite winter types, 124 semi-winter types, and 13 spring types were used for GWAS (Supplementary Table 1). Based on their origins, 112 accessions originated from China, 24 from Oceania, 5 from Europe, 1 from North America, and 1 from India. This GWAS panel also included parental lines; R1 and R2 utilized for the development of DH and IF_2 populations investigated in this study. The seeds of all accessions were procured from the National Mid-term Genebank for Oil Crops, Wuhan, China, and then multiplied at the Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences (OCRI-CAAS), Wuhan, China. All accessions were planted in a field following a randomized complete block design with 2 replications in 3 consecutive years (2011, 2012, and 2013) at Yangluo Research Station (248 310S; 338 00E) in Hubei, China. Seeds were sown at normal agronomic density in plots of 2 × 1 m. Each plot contained three rows; each row with 18 plants. Field management followed the standard agricultural practice.

DH Genetic Mapping Population

A mapping population, designated as RR, comprising 96 DH lines was developed from an F_1 plant derived from the cross of R1 (maternal parent) and R2 (paternal parent). The R1 and R2 were elite semi-winter breeding lines developed by OCRI-CAAS. R1 is a highly resistant advanced breeding line to pod shatter (Liu J. et al., 2013) whereas R2 is a highly prone to pod shattering line under field conditions; both lines are paternal lines of two high yielding commercial hybrid cultivar in China. The RR-DH population was grown in consecutive 2 years, i.e., 2013 and 2014 under winter-cropped environments at Yangluo Research Station and phenotyped for pod shatter resistance.

Construction of Immortalized F_2 (IF_2) Validation Population

In order to verify the genetic associations between SNP markers and pod shatter resistance identified in a RR-DH population and to understand additive interaction among loci, all DH lines were intercrossed following a random permutation design (Hua et al., 2002) for constructing an immortalized F_2 (IF_2) population. The random permutation was repeated three times. In each permutation, the 96 DHs were randomly divided into two groups, and the 48 lines in each group were paired up at random to a counterpart in the other group by taking one line from each

group for one cross at a time and taking one from the rest lines for the next cross to ensure that each DH line was used only once in each round of permutation. Pairs with the same two parental lines from the three repeated permutation were manually corrected to eliminate identical pairings. In theory, 48 IF_2 crosses should be produced from each round and in total 144 crosses could be obtained from the three repeats. However, some combinations failed to obtain seeds due to an asynchronous flowering of the parental DH lines, resulting in a total of 124 IF_2 derivatives. All parental DH lines and their hybrid derivatives (F_1) were planted in a randomized complete block design in Yangluo Experimental Station in 2013 winter season. Seeds were sown at normal agronomic density in plots (2×1 m/plot). Each plot contained three rows with 18 plants in each row. Field management followed the standard agricultural practice.

Assessment for Resistance to Pod Shattering

At physiological maturity, 10 plants from the middle of the plots were harvested to evaluate their resistance to pod shatter. Ten pods from each plant were taken from the main inflorescence and then bulked to make a composite sample for measuring pod shatter resistance index (PSRI) using a modified random impact test (RIT; Peng et al., 2013). Samples of mature pods were first oven dried at 45°C for 8 h and then subjected to shaking at 300 rpm in a drum with an inner diameter of 20 cm and a height of 12 cm, together with ball bearings (14 mm diameter). In this laboratory-based RIT procedure, the number of dispersed pods was recorded five times at 2 min intervals of standardized shaking. The PSRI was calculated as follows: $\text{PSRI} = 1 - \sum_{i=1}^{i=5} xi \times (6 - i) / 100$, where x_i is the number of ruptured pods at the ith time ($1 \leq i \leq 5$).

SNP Genotyping

Genomic DNA was isolated from pooled samples of young leaves from 5 plants of each genotype using a CTAB method (Saghai-Maroof et al., 1984). DNA content of each sample was measured using Nanodrop spectrometer (Model ND-2000). The DNA samples were genotyped with the Illumina Brassica 60K Infinium® SNP array as per manufacture's protocol (Illumina Inc., San Diego, USA) by Emei Tongde Co. (Beijing). The SNP data were clustered and called using the Genome Studio genotyping software (Illumina). Among the three possible genotypes (AA, AB, and BB), genotypes with AB alleles was excluded, the remaining homozygous SNP markers were selected to carry out genetic analyses. Genotypic data were curated to remove those SNPs with AA or BB frequency equal to zero, call rates $\geq$0.8 and minor allele frequency $<$0.05.

Construction of a High Density SNP Genetic Map

The software IciMapping V4.0 (Wang et al., 2014, http://www.isbreeding.net/software/?type=detail&id=14) was used to "bin" redundant markers with exactly the same genotypes. Distortion in segregating SNP markers was checked using the χ^2 test according to the expected segregation ratio [AA(1): BB(1)] in DH population. Non-redundant SNP markers showing 1:1 segregation ratio were then used for construction of the genetic linkage map using the software JoinMap version 4.0 (Stam, 1993, https://www.kyazma.nl/index.php/mc.JoinMap), using a recombination frequency of $<$0.25 and minimum LOD score of 5. Recombination frequencies were converted using Kosambi's algorithm (Kosambi, 1944). Linkage groups were assigned to chromosomes A01 to A10 and C01 to C09 according to published genetic maps (Liu L. et al., 2013; Brown et al., 2014; Wang et al., 2015).

In silico Mapping of SNP Markers

In order to verify the chromosomal location of SNP markers and to compare their physical positions in relation to the known genes involved in pod shatter resistance in *Arabidopsis thaliana* and *B. napus* (www.tair.com, Girin et al., 2010; Hu et al., 2012; Raman et al., 2014; Dong and Wang, 2015), sequences of all associated SNPs and candidate genes were used to perform BlastN searches against the *B. napus* cv. Darmor genome sequence (Chalhoub et al., 2014). Only the top blast-hits with an *E*-value cut-off of $1E^{-15}$ were considered for genetic and comparative analyses. The closest known pod shatter resistance gene in relation to the physical position of SNP marker on the *B. napus* genome was assumed to be a "candidate" gene for pod shatter resistance in genetic mapping populations.

Statistical Analysis and QTL Identification

The PROC GLM procedure was used to estimate the variance components for individual traits/environments using SAS software version 8.1 (SAS Institute Inc., 1999). Genotype was considered a fixed effect, whereas environment was considered as random effects. The mean value of the trait was calculated and then used for genetic analysis.

The model of composite interval mapping (CIM) in the WinQTL cartographer version 2.5 (Wang et al., 2007) was used for QTL identification. Multiple linear regression was conducted using forward-backward stepwise and a probability model was set with 0.05 and window size at 10 cM. The LOD threshold was determined by 1000 permutation test (Churchill and Diverge, 1994) and a significant level of 0.01 were selected to determine whether there is any QTL for pod shatter resistance.

Population Structure, Kinship, and GWA Analysis

For GWAS, three data types are required: genotypic data, population structure within the GWA panel (population) and phenotypic trait information. After discarding SNP markers which were either monomorphic and/or had minor allele frequencies (MAFs) $<$0.05, a total of 66.1% (34,469/52,157) high-quality polymorphic SNPs were selected for GWAS.

In order to infer the population structure of the GWAS panel, a subset of data of 2434 SNP markers which showed genome-wide coverage across all 19 chromosomes were used into the software package STRUCTURE version 2.3.4 (Pritchard et al., 2000). An admixture model was performed for five independent runs with a *K*-value, ranging from 1 to 10, iterations of 100,000 times, burn-in period of 100,000 MCMC (Markov Chain Monte Carlo).

The optimal *K*-value was determined according to the method of Evanno et al. (2005). The cluster membership coefficient matrices of replicate runs from STRUCTURE were integrated to get a Q matrix by the CLUMPP software (Jakobsson and Rosenberg, 2007). Accessions with the probability of membership >0.7 were assigned to corresponding clusters, and those <0.7 were assigned to a mixed group. Q matrices were used as covariates to calculate population structure with K. The extent of LD for each chromosome was estimated using pairwise r^2 of all mapped SNPs using window of 500.

With Best linear unbiased predictors (BLUPs) of calculated for all phenotypic environments (3 years, **Table 1**), we conducted a GWAS with 34,469 genome-wide SNPs using a univariate unified mixed linear model (Yu et al., 2006) that eliminated the need to recomputed variance components (i.e., population parameters previously determined, or P3D; Zhang et al., 2010). To control the effect of familial relatedness in GWAS, the kinship matrix based on coancestry (Loiselle et al., 1995) was estimated using 34,469 genome-wide SNPs. A likelihood-ratio-based R^2 statistic, denoted R^2 LR (Sun et al., 2010), was used to assess the amount of phenotypic variation explained by the model. The Benjamini and Hochberg (1995) procedure was used to control the multiple testing problem at false-discovery rates (FDRs) of 5 and 10%. GWAS was performed by TASSEL 4.0 (Bradbury et al., 2007) using a mixed linear model (MLM) in which relative kinship matrix (K) and population structure (Q) were included as fixed and random effects, respectively. Significance of associations between traits and SNPs was set on threshold $P < 2.90 \times 10^{-5}$ (i.e., $-\log_{10}(p) = 4.5$). The threshold is 2.90×10^{-5} at a significant level of 1% after Bonferroni multiple test correction (1/34,496). Furthermore, the false discovery rate (FDR at $P < 0.05$) was applied to estimate the proportion of false positives among the significant associations (Dabney and Storey, 2004). The marker effect and the significant value generated in R package for each SNP were exported (http://cran.r-project.org). LD block analysis was performed as described previously, keeping the lead SNP within each LD block (Gabriel et al., 2002).

Allelic Effects of Pod Shatter Accessions

Based on pod shatter resistance indices, all 143 accessions were ranked and then investigated for allelic diversity at significant GWAS SNP loci. PSRI of R1 and R2 were 0.45 and 0.04, respectively. Accessions having PSRI $\geq$ 0.28 were assumed to have superior alleles for pod shatter resistance.

RESULT

Genetic Variation for Pod Shatter Resistance in Biparental Populations

Predicted means for PSRI of DH and IF_2 populations showed a continuous distribution for pod shatter resistance irrespective of growing environments. Both parental lines differed significantly in pod shatter resistance across all phenotyping environments. R1, the resistant parent, had consistently higher PSRI (0.45) compared to the pod shatter prone parent, R2 (0.04; **Figure 1**). The frequency distribution of PSRI deviated significantly from normality among DH and IF_2 lines ($P < 0.001$). Among RR-DH lines, a strong positive correlation ($r = 0.60$) of genotype performance for PSRI was observed across 2013 and 2014 environments (**Figure 2**), suggesting that phenotypic variation in PSRI is genetically controlled, consistent with high broad-sense heritability values (**Table 1**). Analysis of variance showed that the effects of genotype (G), and genotype × environment (G × E) interaction on PSRI were significant (**Table 1**), suggesting that genetic mapping populations must be evaluated across multiple sites/years to ensure valid phenotypic assessment.

Construction of a High-Density Genetic Bin Map for QTL Analysis

Of the 52,157 SNP markers (60K Infinium array), only 16.4% (8540) were polymorphic between the parental lines, R1 and R2 of the RR-DH population. Of these, 7804 SNP markers showing 1:1 segregation ratio, as determined by the χ^2 test ($P = 0.05$), were used for construction of a genetic linkage map and QTL analysis. A majority (99%) of the polymorphic markers (7728/7804) were anchored to the 19 chromosomes of *B. napus* and mapped to 2046 distinct loci, with 1384 loci on A genome, and 662 loci on the C genome (**Table 2**, **Figure 3**). A total of 5682 SNP loci showed co-segregation and could be grouped into 900 discrete bins. A genetic linkage map of RR-DH population spanned 2217.2 cM of Kosambi map distance. The marker density of the 19 chromosomes ranged from 0.61 (A03) to 2.96 (C09), with an average of 1.08 cM. The chromosome A03 displayed the maximum marker density (738 markers representing 222 loci) and chromosome C09 had the least density (77 markers representing 24 loci). In particular, chromosomes C08 and C09 were shorter (66.6–71 cM) than rest of the chromosomes (**Table 2**).

The SNP genotypes of 124 F_1 hybrids were deduced from their corresponding DH parental lines to provide a bin map for the IF_2 crosses (**Figure 3**). There were three genotypes in each bin: homozygous genotype from R1 (*MM*), homozygous genotype from R2 (*mm*), and heterozygous genotype (*Mm*). The average proportion of three genotypes for each cross was 27.3, 29.2, and 43.5%, respectively. Therefore, the composition of genotypes in IF_2 was similar to that in an F_2 population. This population could therefore be used to detect QTL with the same analytical method used for an F_2 population.

QTL Associated with Pod Shattering Resistance in a RR-DH Population

In the RR-DH population, four significant QTL *qSRI.A01a*, *qSRI.A06a*, *qSRI.A06b,* and *qSRI.A09* were detected for PSRI on chromosomes A01, A06, and A09 (**Table 3**). These QTL accounted for 5.66–16.91% of the phenotypic variation. The *qSRI.A09* (LOD = 4.31–7.69) accounted for the maximum phenotypic variation in pod shatter resistance (9.81–16.9%). Two QTL, *qSRI.A09* delimited with the SNP Bn-A09-p30171993 (A09) and *qSRI.A06b* delimited with the SNP marker Bn-A06-p115948 (A06) were repeatedly detected across both

TABLE 1 | Genetic variation and broad-sense heritability in pod shatter resistance index (PSRI) among three populations.

Population	Phenotyping environment (year)	PSRI (range)	PSRI (Mean ± SD)	CV (%)	Genotype (G)	Environment (E)	G × E	H^2(%)
DH	2013	0.01–0.58	0.14 ± 0.12	92.49	**	**	**	85.11
	2014	0.03–0.99	0.46 ± 0.28	61.48				
IF_2	2014	0.05–0.99	0.50 ± 0.27	53.70	**			
GWAS diversity set	2011	0.00–0.58	0.09 ± 0.11	119.32	**			92.11
	2012	0.01–0.64	0.15 ± 0.14	87.54				
	2013	0.00–0.71	0.09 ± 0.14	154.04				

***P < 0.01 for the effect of genotype (G), environment (E), and genotype by environment interaction (G × E) on phenotypic variance estimated by two-way ANOVA.*
CV, coefficient of variation; H^2, broad sense heritability; SD, Standard deviation.

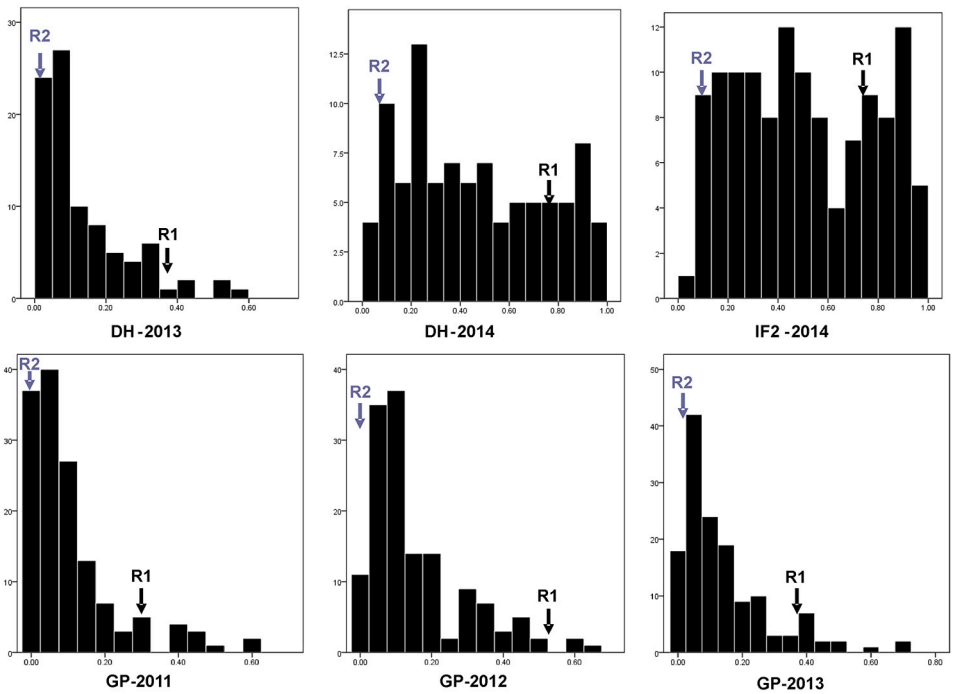


FIGURE 1 | Phenotypic distribution for individual pod shatter resistance index in DH population (R1 × R2) across 2 years (2013 and 2014) and IF_2 population (2014), and GWAS panel (GP) across 3 years (2011, 2012, and 2013). Transgressive segregation was observed in the DH and IF_2 populations from all of the environments. Gray arrows are for R2 and black arrows are for R1. y axis represents Number of lines and accessions and x axis represents pod shatter resistance index measured by RIT (Random Impact Test) method.

environments in 2013 and 2014. It is possible that QTL *qSRI.A06a* and *qSRI.A06b* may be the same, as both were detected in close proximity of Bn-A06-p15913910/Bn-A06-p115948 markers, mapped within 250 kb on the physical map of *B. napus* genome (**Table 3**, Supplementary Table 4). The pod shatter resistant parent, R1 contributed favorable alleles for pod shatter resistance based on RTI at all QTL detected (**Table 3**), consistent with the high pod shatter resistance index of R1 compared to R2 (**Figure 1**).

Verification of Loci Associated with Pod Shatter Resistance in IF_2 Population

In order to verify the allelic effects of QTL revealed in a RR-DH population (**Table 3**), we performed an independent linkage analysis for association between SNP markers and genetic variation in pod shatter index evaluated in an IF_2 population (**Figure 1**, **Table 3**). We identified four QTL, *qSRI.A01b*, *qSRI.A03*, *qSRI.A06b,* and *qSRI.A09* for PSRI on chromosomes A01, A03, A06, and A09, respectively (**Table 3**). Two consistent and stable QTL *qSRI.A06b and qSRI.A09,* as identified in RR-DH population, were also detected in an IF_2 population. The same set of markers, Bn-A06-p115948 (A06) and Bn-A09-p30171993 (A09) revealed significant phenotypic variation for pod shatter resistance (**Table 3**). Significant QTL, *qSRI.A01b* (A01) and *qSRI.A03* (A03) were defined by the SNP markers Bn-A01-p2365493 and Bn-scaff-22728-1-p75030, respectively (**Table 3**). These QTL accounted for up to 13.14% of phenotypic variation in PSRI.

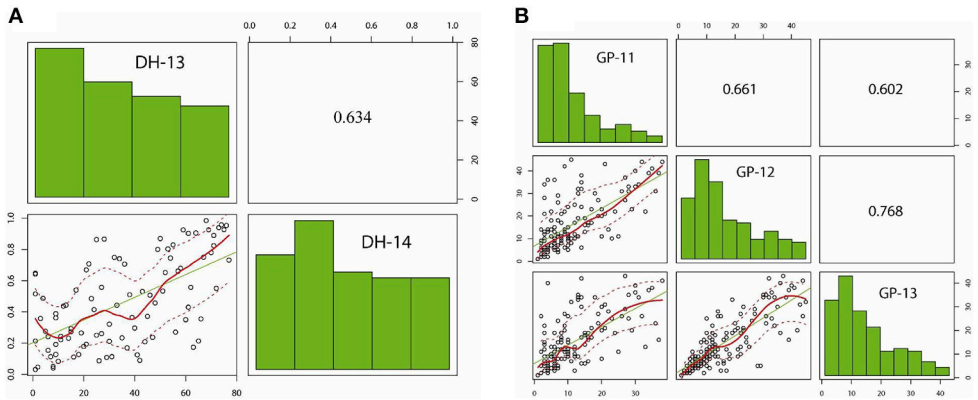


FIGURE 2 | Distribution of pod shatter resistance, as measured with the random impact test, among DH lines from the R1/R2 and GWAS diversity set. Pair-plots of EBLUPS from DH lines and GWAS diversity set showing correlations are presented. **(A)** R1/R2 population grown under two environments: experiment 1 (DH-13); experiment 2 (DH-14). **(B)** GWAS diversity set grown under three environments: GP-11, GP-12, and GP-13.

TABLE 2 | Features of the genetic linkage map of a DH population derived from R1 /R2 of *B. napus.**

Chromosome	No. of polymorphic markers	No. of mapped markers	No. of mapped markers in bin loci	No. of bin loci	Map length (cM)	Average distance between loci (cM)
A01	329	233	145	49	95.9	0.66
A02	317	269	98	50	117.5	1.20
A03	738	612	222	96	136.5	0.61
A04	191	448	139	60	126.8	0.91
A05	615	498	199	84	125.1	0.63
A06	450	359	158	67	133.2	0.84
A07	483	417	122	56	98.8	0.81
A08	269	237	52	20	108.2	2.08
A09	632	546	147	61	135.1	0.92
A10	191	115	102	26	118.9	1.17
C01	405	361	81	37	147.9	1.83
C02	891	852	94	55	104.4	1.11
C03	665	602	123	60	134.3	1.09
C04	574	526	106	58	163.4	1.54
C05	147	116	56	25	123.8	2.21
C06	500	469	73	42	110.1	1.51
C07	794	769	39	14	99.7	2.56
C08	272	239	66	33	66.6	1.01
C09	77	60	24	7	71.0	2.96
Subtotal for the A genome	4215	3734	1384	569	1196.0	0.86
Subtotal for the C genome	4325	3994	662	331	1021.2	1.54
Total (A+C)	8540	7728	2046	900	2217.2	1.08

**Markers which showed co-segregation with each other were binned using the ICI mapping package (http://www.isbreeding.net/software/?type=detail&id=14).*

GWAS Analysis for Pod Shatter Resistance in a Diversity Panel

In order to identify loci associated with pod shatter resistance in a diverse panel of accessions, exploiting the historic recombination events, we conducted a GWAS using the Q + K model accounting both for population structure as well as kinship relatedness (Bradbury et al., 2007). Based on a probability-of-membership (a measure of population structure) with threshold of 70%, a diversity panel of 143 lines could be assigned to three groups (group I: 17 lines, group II: 99 lines, and group III: 27 lines representing a mixed group; Supplementary Table 1). In addition, cluster analysis was

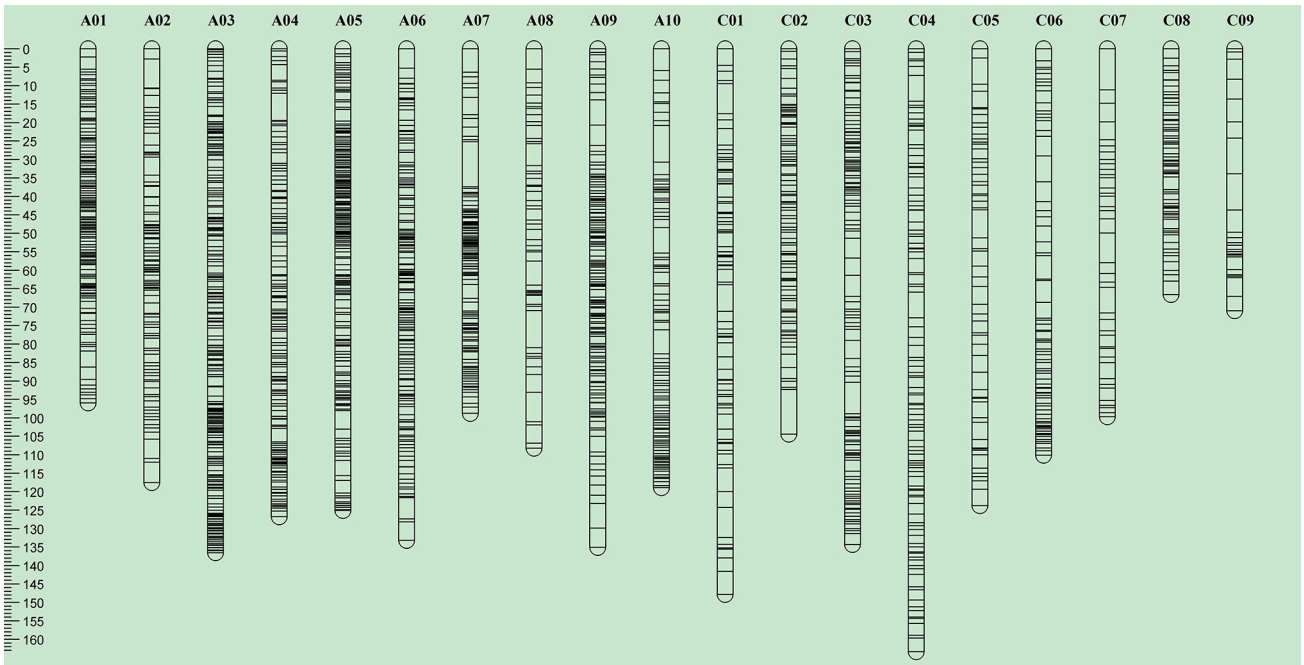


FIGURE 3 | Overview of genome-wide SNP density in the bin map of the RR-DH population derived from R1 (resistant to pod shatter) and R2 (prone to pod shatter) lines of *B. napus*. The ordinate shows the genetic distance along each of the 19 linkage groups corresponding to the 19 *B. napus* chromosomes. Scale in Kosambi centimorgans (cM) is on the left.

conducted; the Neighbor-Joining phylogenetic tree based on Nei's genetic distances displayed two clear clades (Supplementary Figure 1), reconfirming the presence of two groups (group I and II, Supplementary Table 1) estimated by STRUCTURE. Estimates of an average nucleotide diversity (also known as polymorphism information content or PIC) of 0.366 showed that the overall genetic variation in the germplasms studied here represents ∼62.9% of the rapeseed diversity (PIC > 0.35; Supplementary Table 2). In order to test the robustness of population structure revealed by cluster analysis, we also used the Δk method (Evanno et al., 2005). The 143 accessions could be divided into two sub-populations (Supplementary Figure 2). The average relative kinship between any two lines was 0.0332, or ∼57% of the pairwise kinship estimates were close to 0, and 21% of the kinship estimates ranged from 0 to 0.05 (Supplementary Figure 3). The genome-wide LD decay of each chromosome for rapeseed germplasms is shown in Supplementary Figure 4.

GWAS detected a total of 38 SNPs that showed significant association (up to $P < 2.90E^{-5}$) with pod shatter resistance across three environments (**Table 3**, Supplementary Table 3). After Bonferroni correction, we identified 6 genomic regions (QTL) on chromosomes A01, A06, A07, A09, C02, and C05 accounting for up to 45.9% cumulative phenotypic variance for pod shatter resistance in a GWAS panel (**Table 3**). Multiple environment analyses revealed that at least two QTL, *qSRI.A06b* delimited with the SNP marker Bn-A06-p115948 (A06) and *qSRI.A09* delimited with the SNP Bn-A09-p30171993 (A09) could be repeatedly detected across populations (DH, IF_2, and GWA panel) as shown in **Table 3**. Significant QTL associated with SNPs Bn-A07-p7392457 (A07), Bn-scaff_15712_6-p214229 (C02), and Bn-scaff_17869_1-p1058624 (C05) were not detected in both RR-DH/IF_2 genetic mapping populations.

Physical Mapping of Significant QTL for Pod Shatter Resistance in Comparison to Previously Detected QTL and Candidate Genes

In order to gain insights of genetic architecture of pod shatter resistance loci, we compared the physical positions of markers associated with QTL identified in this current and previously studies (Hu et al., 2012; Raman et al., 2014). The sequences of markers significantly associated with pod shatter resistance were subjected to BLAST against the physical reference genome of *B. napus*. The markers linked with pod shatter resistance loci on chromosome A09: NS380 and NS381 (Hu et al., 2012), DArTseq markers 3146978 and 3105723 (Raman et al., 2014) and Bn-A09-p30171993 (this study) were located within ∼400 kb region of *B. napus* genome (**Figure 4**). This genomic region delimited from 30.84 to 31.98 Mb of *B. napus* genome also contains QTL having major allelic effects for pod length and seed weight in rapeseed (Li et al., 2014; Fu et al., 2015). A recent research showed that the *AUXIN RESPONSE FACTOR 18* (*ARF18*) gene affecting seed weight and pod length is located within this region (Liu et al., 2015). These studies suggested that *qSRI.A09* is a hotspot region for seed yield and pod traits such as pod shatter resistance and pod length in rapeseed. The major QTL genomic regions on A09 (**Table 3**) were consistent as

TABLE 3 | Comparison of QTL identified for pod shatter resistance from linkage mapping and association analysis of mapping populations.

Mapping population	Phenotyping year	QTL	Chr[a]	LOD	*P*-Value for GWAS association	Lead SNP with highest *P*-value	Physical map position	Genetic map position (cM)*	Confident Interval	R^2 (%)	Candidate gene
DH	2014	*qSRI.A01a*	A01	5.43		Bn-A01-p11702957	11221208	71.5	69.5–75.4	13.14	
IF_2	2014	*qSRI.A01b*	A01	7.07		Bn-A01-p2365493	1858394	56.9	54.3–61.5	12.92	*SPATULA (SPT)*
GWAS		*qSRI.A01c*	A01		8E-06	Bn-A01-p10523833	883265	–	–	6.37	
IF_2	2014	*qSRI.A03*	A03	3.84		Bn-scaff_22728_1-p75030	5375993	62	58.7–64.3	4.01	
DH	2013	*qSRI.A06a*	A06	2.93		Bn-A06-p15913910	17373387	60.4	56.2–63.4	5.66	*GIBBERELLIN 3-OXIDASE 1*
DH	2014	***qSRI.A06b***	**A06**	2.88		Bn-A06-p115948	79870	57.8	54–59.8	6.47	*GIBBERELLIN 3-OXIDASE 1*
IF_2	2014	***qSRI.A06b***	**A06**	4.19		Bn-A06-p115948	79870	58.3	55.1–62.3	3.69	
GWAS		***qSRI.A06b***	**A06**		1.5E-05	Bn-A06-p115948	79870	–	–	6.61	
GWAS		*qSRI.A07*	A07		1E-06	Bn-A07-p7392457	858774	–	–	7.3	*YABBY1*
DH	2013	***qSRI.A09***	**A09**	7.69		Bn-A09-p30171993	3297223	68.1	65.8–75.9	16.91	*SHATTERPROOF 1/2, ARF18*
DH	2014	***qSRI.A09***	**A09**	4.31		Bn-A09-p30171993	3297223	68.1	67.4–76.5	9.81	
IF_2	2014	***qSRI.A09***	**A09**	12.41		Bn-A09-p30171993	3297223	68.1	67.4–76.5	10.89	
GWAS		***qSRI.A09***	**A09**		4.4E-09	Bn-A09-p30171993	3297223	–	–	12.07	
GWAS		*qSRI.C02*	C02		6.5E-06	Bn-scaff_15712_6-p214229	40565480	–	–	6.32	
GWAS		*qSRI.C05*	C05		1.1E-06	Bn-scaff_17869_1-p1058624	19589640	–	–	7.23	

[a]*Chromosome;* * *not applicable.*

Consistent QTL identified across mapping populations/environments are in bold.

reported previously, suggesting that indeed QTL identified herein are relevant to international germplasm and rapeseed breeding programs.

In order to identify putative candidate genes involved in pod shatter resistance in the mapping populations (GWAS, DH, and IF_2) investigated herein, we compared the physical map positions of SNP markers that showed significant associations in GWAS and mapping populations and known candidate genes involved in positively and negatively regulation of pod shatter such as *FILAMENTOUS FLOWER*, *YABBY3*, *ASYMETERICAL LEAVES1/2*, *BREVIPEDICELLUS*, *SHATTERPROOF1/2*, *INDEHISCENT*, *ALCATRAZ*, *FRUITFUL*, *APETELA2*, *NAC SECONDARY WALL THICKENING PROMOTING FACTOR1*, *SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1*, *DEHISCENCE ZONE POLYGALACTURONASE1*, *SPATULA*, and *PIN3* (reviewed in Dong and Wang, 2015) (Supplementary Table 4) on the sequenced *B. napus* genome. Among these significant SNPs underlying genetic variation for pod shatter resistance (**Table 3**), Bn-A01-p2365493 at the *qSRI.A01b* (A01) was mapped to a candidate gene *SPATULA;* Bn-A06-p15913910 and Bn-A06-p115948 corresponding with *qSRI.A06a* (A06) and the *qSRI.A06b* (A06) were all mapped to candidate genes *GIBBERELLIN 3-OXIDASE 1;* Bn-A07-p7392457 at the *qSRI.A07* (A07) was mapped to a candidate gene *YABBY1*. Except that, Bn-A09-p30171993 at the *qSRI.A09* (A09) was mapped two homologous regions on A09 and C08 which is within 11 kb from the *ARF18* gene controlling seed weight and pod length in *B. napus* (Liu et al., 2015). Both copies of *ARF18* in *B. napus*; *BnaA.ARF18.a* and *BnaA.ARF18.c* were also located on the physical positions of chromosomes A09 and C08, respectively (Supplementary Table 4, **Figure 4**). PCR marker, Shp-100925 associated with *BnSHP-1* locus on chromosome A09 was also mapped in the vicinity of *qSRI.A09* and *ARF18* (**Figure 4**).

Allelic Diversity at Significant QTL Associated with Pod Shatter Resistance

Based on the PSRI ranking of 143 accessions used for GWAS, 18 elite cultivars having PSRI $\geq$ 0.28 were selected and their allele diversity was investigated at QTL *qSRI.A01*, *qSRI.A06b*, *qSRI.A07*, *qSRI.A09*, *qSRI.C02,* and *qSRI.C05* that showed significant associations with lead SNP markers (**Tables 3, 4**). These 18 accessions were originated from 5 provinces of China (Supplementary Table 1), representing the main rapeseed production area of the Yangtze River eco-region. About one-half of the resistant accessions, including the top five with PSRI $\geq$ 0.44 (**Table 4**), all originated from Hubei province in the middle Yangtze River eco-region, shared the "CC" SNP allele at Bn-A09-p30171993 locus. Generally, the resistant accessions possess multiple favorable alleles suggesting the potential for recombining them in a breeding design to improve resistance to pod shatter in rapeseed breeding programs. For example, the most resistant genotype, Zhongshuang2 might be further improved through complementary recombination with the favorable alleles (CC) of Bn-A09-p30171993 from other resistant accessions identified in this study (**Table 4**). In

addition, combining favorable alleles among other accessions would also improve pod shatter resistance within a breeding program.

DISCUSSION

Genetic Variation for Pod Shatter Resistance in Rapeseed

In this study, we determined the extent of genetic variation for pod shatter resistance in bi-parental DH and IF_2 populations, and GWAS diversity panel comprising 143 accessions representing released Chinese cultivars/advanced breeding lines. We identified seven accessions with PSRI $\geq$ 0.4 across years which exhibited improved levels of PSRI such as Zhongshuang2, OG3151, and Zhen2609, compared to standard check cultivars and would provide valuable resources for genetic improvement of pod shatter resistance in rapeseed improvement programs. However, we could not benchmark the level of resistance to pod shatter among accessions utilized in this study and previous ones (Wen et al., 2008; Pu et al., 2013, Raman et al., 2014), due to different assessment methods, germplasm, and growing conditions. Previous studies showed that there is a limited natural variation for pod shatter resistance in rapeseed (Wen et al., 2008; Raman et al., 2014), which has contributed to the lack of significant genetic improvement for this trait in breeding programs. It is possible that improved pod shatter resistance characterized herein may have been derived from pod shatter resistant sources of *B. rapa*, as they have been extensively used for introgression of novel alleles for traits of interest as well as to expand genetic base of rapeseed germplasm especially in China (Qian et al., 2005; Zou et al., 2010). Sources of pod shatter resistance are well documented in *B. rapa* gene pool and have been exploited in breeding programs (Kadkol et al., 1985, 1986; Mongkolporn et al., 2003; Hossain et al., 2011; Raman et al., 2014).

A laboratory based method (RIT) proved to be robust in determining the extent of pod-shatter resistance across several experiments. Further research efforts are needed to validate RIT for pod shatter resistance with pendulum test and field based methods such as delayed harvest across rapeseed growing regions.

Genetic Basis of Phenotypic Variation in Pod Shatter Resistance

We utilized both classical QTL and GWAS approaches to detect genomic regions associated with pod shatter resistance (**Table 3**). Both these approaches have their own advantages and disadvantages in QTL detection. For example, classical linkage analysis has strong statistical power and proven to be effective in detecting QTL, but only capture the recombination events in two parents used in constructing bi-parental DH/intercross populations. GWA simultaneously detects multiple alleles at the same locus, due to the accumulation of historical recombination events during systematic selection in breeding and resolves QTL based on LD particularly in species such as rapeseed where LD decays rapidly (Flint-Garcia et al., 2003; Buckler et al., 2009; Gajardo et al., 2015). The combined application of both approaches; QTL and GWAS not only improve the efficiency of QTL detection, but also facilitate the identification of reliable and stable QTL and novel alleles across a wide range of germplasm (Krill et al., 2010; Raman et al., 2014, 2016).

In this study, we identified six QTL associated with pod shatter resistance which accounted for up to 50% the phenotypic variation in PSRI in DH and IF_2 mapping populations. Previously, several QTL associated with pod shatter resistance were identified in a DH mapping populations derived from ZY72360/R1, H155/Qva, and BLN2762/Surpass400, and in diverse panel of accessions of *B. napus*, originated from Australia, China, and Europe (Hu et al., 2012; Wen et al., 2013; Raman et al., 2014). For example, Wen et al. (2013) identified 13 QTL for pod shatter resistance on the chromosomes A01, A04, A07, A08, C05, and C08; however only three of them were consistent at both locations. Recently, Raman et al. (2014) identified 12 QTL associated with pod shatter resistance in a DH population from BLN2762/Surpass400 on chromosomes A03, A07, A09, C03, C04, C06, and C08 using DArTseq markers. *In silico* mapping analysis of Illumina SNP markers showed that some of the QTL identified in this study are similar as reported previously (Raman et al., 2014) such as on A01, A03, and A09. Two QTL *qSRI.A06* (A06) and *qSRI.A09* (A09) were detected repeatedly across DH and GWAS populations and phenotypic environments, implicating their involvement in pod shatter resistance in rapeseed cultivars of Chinese origin. This suggests that there were at least two genes involved in resistance to pod shattering in DH and IF_2 populations derived from R1. In a previous study (Hu et al., 2012), one major quantitative trait locus *psr1* on chromosome A09 accounting 47% of phenotypic variation in pod shatter resistance was identified in an F_2 population derived from ZY72360/R1. Comparative analysis of the A09 locus in the linkage maps of BLN2762/Surpass400 (Raman et al., 2014) and R1/R2 (this study) with *B. napus* physical map, showed an inversion event of the 400 kb QTL interval *qSRI.A09*/*Qrps.wwai-A09*. This result is partly consistent with the previous comparative genomic studies showing rearrangements in the A subgenome of *B. napus* (Xu et al., 2010; Li et al., 2014).

The present study showed that the PCR marker, Shp-100925 associated with *BnSHP-1* locus was mapped in the vicinity of *qSRI.A09* and *ARF18* (**Figure 4**). The role of auxin in pod dehiscence and other developmental processes has been documented in Arabidopsis (Okushima et al., 2005; Sorefan et al., 2009), *B. juncea*, and *B. napus* (Jaradat et al., 2014). For example, Sorefan et al. (2009) reported that a local auxin minimum is required for the formation of valve margin separation layer for seed dehiscence which is controlled by *IND* gene. *ARF18* gene also regulates cell growth in the pod wall via auxin-response pathway in *B. napus* and simultaneously affects seed weight and pod length in an F_2 population derived from the ZY72360/R1 (Liu et al., 2015). In a recent study, auxin biosynthesis, transport, and signaling was shown to be repressed in *B. juncea* (less prone to shattering) compared to *B. napus* (more prone to pod shattering) genotypes (Jaradat et al., 2014). These studies suggest that that the auxin minimum may be responsible for pod shatter trait in the mapping populations investigated here.

TABLE 4 | SNP alleles at the significant QTL identified for pod shatter resistance in DH, $IF_{2,}$ and GWAS panel among 18 pod shatter resistant accessions.

Name	SRI mean	A01		A06		A07		A09		C02		C05	
		Bn-A01-p10523833 A/G		Bn-A06-p115948 A/G		Bn-A07-p7392457 T/G		Bn-A09-p30171993 A/C		Bn-scaff_15712_6-p214229 A/G		Bn-scaff_17869_1-p1058624 T/G	
Zhongshuang2	0.49			AA		TT		AA		AA			GG
OG3151	0.47	AA			GG		GG		CC	AA		TT	
Zhen2609	0.46			AA					CC	AA			GG
R1	0.45			AA		TT			CC	AA			GG
OG3237	0.44	AA			GG		GG		CC	AA		TT	
9905	0.43		GG	AA					CC		GG		GG
OG3190	0.41				GG					AA			GG
HX0352	0.39						GG				GG	TT	
L229	0.38			AA			GG				GG		GG
1055B	0.35			AA			GG	AA		AA			GG
9490	0.35		GG	AA			GG		CC	AA			GG
9230	0.34	AA		AA			GG		CC	AA			GG
9226	0.31		GG	AA			GG		CC	AA			GG
Zhongshuang11	0.31			AA		TT			CC	AA			GG
L233	0.29			AA			GG			AA		TT	
Huyou19	0.29				GG				CC	AA		TT	
Fan189	0.28				GG			AC		AA			GG
OG3186	0.28				GG					AA			GG
R2	0.04	AA			GG		GG	AA			GG	TT	

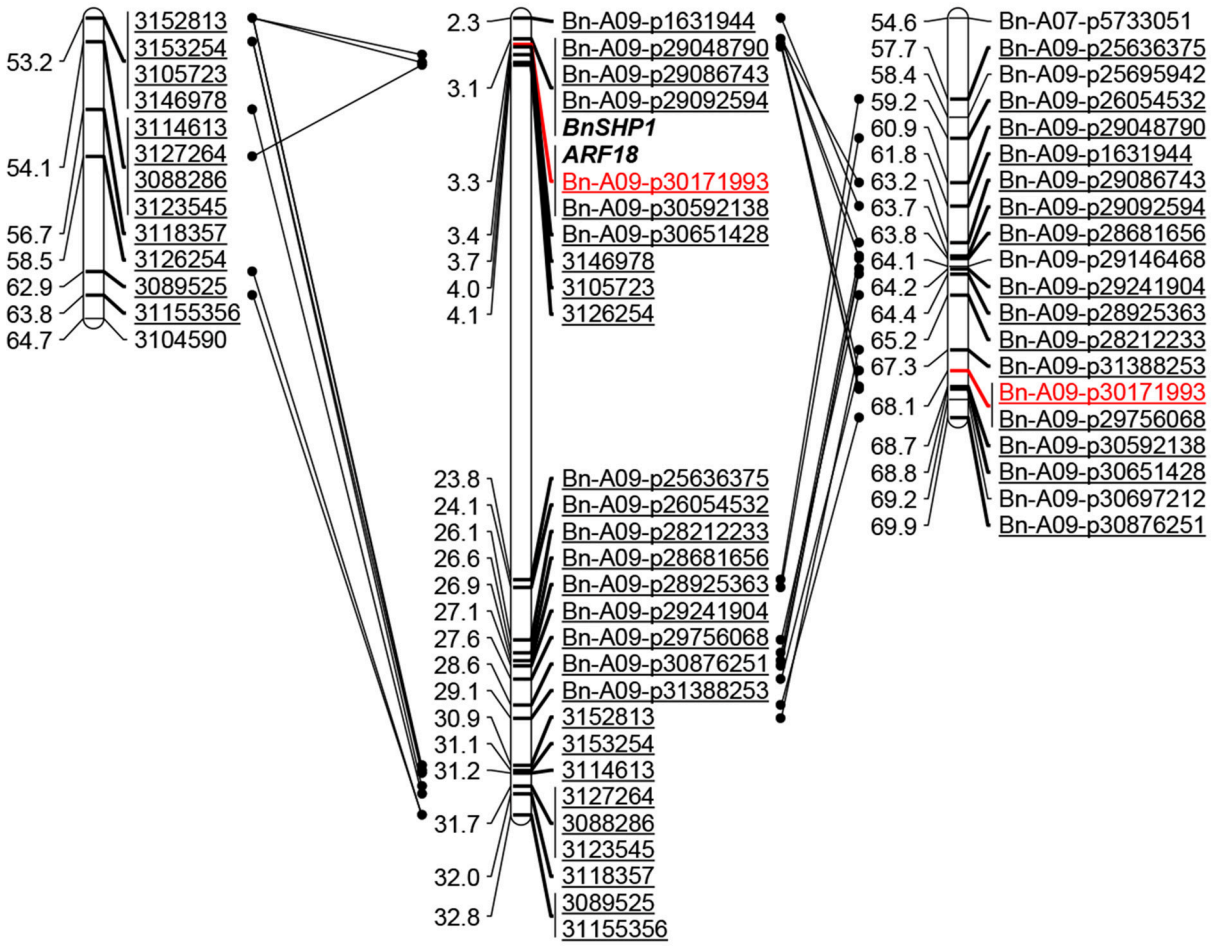


FIGURE 4 | Comparative analysis of *qSRI.A09* on chromosome A09 for pod shatter resistance in the RR-DH and previously mapped populations. Left: the DArTseq markers in the order of the genetic map (cM) for *B. napus* based on a previous study (Raman et al., 2014). **Middle**: the markers in the order of the physical map (Kb) for *B. napus* (Darmor-bzh). Physical map distances are given in fraction (1/1,000,000th) of the actual coordinates of the *B. napus* genome. The markers in red are the most associated marker for pod shattering resistance. **Right**: the markers in the order of the genetic map (cM) for R1/R2 (RR-DH) population used in the current study. The marker in red showed highly significant association with pod shatter resistance at *qSRI.A09*.

Further studies are required to establish the role of auxins in genetic variation for pod shattering resistance in diverse *B. napus* accessions.

In addition to *qSRI.A09*/*Qrps.wwai-A09*/*psr1* locus on A09 (Hu et al., 2012; Raman et al., 2014, this study), other QTL *qSRI.A01* (A01), *qSRI.A03* (A03), *qSRI.A07* (A07), *qSRI.C02* (C02), and *qSRI.C05* (C05) also account genetic variation for pod shatter resistance derived from R1, a pod shatter resistant Chinese cultivar. Arabidopsis genes underlying the significant QTL such as *SPATULA* and *GIBBERELLIN 3-OXIDASE 1* (**Table 3**) are likely candidate genes for pod shatter resistance in mapping populations. A basic-helix-loop-helix transcription factor, *SPATULA* is implicated in dehiscence zone in Arabidopsis and regulated by *ARF* (Heisler et al., 2001), suggesting its role in auxin-mediated dehiscence zone formation implicated in pod shatter. GA3ox1 encodes a Gibberellin 3-oxidase, which is a direct and necessary target of *IND* gene (Arnaud et al., 2010). Identification of closely linked markers and the genomic location of QTL on chromosomes A01, A06, A07 and A09 with respect to a reference genome of *B. napus* and the described genes involved in pod shatter resistance of Arabidopsis could also pave the way for map-based cloning of those QTL and unravel the molecular architecture of pod shatter resistance genes in natural germplasm of *B. napus*.

CONCLUSION

Both GWAS and linkage analyses enabled to untangle multiple quantitative trait loci associated with pod shatter resistance in Chinese germplasm of rapeseed. Identification of the improved sources for pod shatter resistance, and understanding the genetic

basis underlying genetic variation in pod shattering resistance in rapeseed germplasm will provide insights into the complex architecture and evolution of this trait which has been subjected to artificial selection since its domestication. SNP markers flanking QTL regions would provide an efficient method for selection of alleles associated with pod shatter resistance in rapeseed breeding programs.

AUTHOR CONTRIBUTIONS

JL and QH conceived and designed the study. JW and HW conducted the DH and IF_2 population experiments; JL and WW carried out the association population experiments; JW and HW analyzed the DH and IF_2 data; JL and WW analyzed the association data; DM and JW produced the DH and IF_2 populations; RZ, HC, and JY did the phenotype assessment; JL, JW, and HR interpreted the data and prepared the manuscript; HR performed comparative and *in silico* analysis; QH supervised the whole study; all authors reviewed and edited the manuscript.

ACKNOWLEDGMENT

We would like to thank the reviewers for their valuable suggestions and comments on this manuscript. The National Key Basic Research Program (2011CB1093), the Natural Science Foundation of China (31471535), the Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (Group No. 118), the Earmarked Fund for China Agriculture Research System (CARS-13), the Hubei Agricultural Science and Technology Innovation Center (201620000001048) and Australian Grains Research and Development Corporation (DAN00208) supported this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01058

REFERENCES

Arnaud, N., Girin, T., Sorefan, K., Fuentes, S., Wood, T. A., Lawrenson, T., et al. (2010). Gibberellins control fruit patterning in *Arabidopsis thaliana*. *Genes Dev.* 24, 2127–2132. doi: 10.1101/gad.593410

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., et al. (2011). Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nat. Biotechnol.* 29, 762–766. doi: 10.1038/nbt.1926

Bayer, P., Ruperao, P., Mason, A., Stiller, J., Chan, C. K., Hayashi, S., et al. (2015). High-resolution skim genotyping by sequencing reveals the distribution of crossovers and gene conversions in *Cicer arietinum* and *Brassica napus*. *Theor. Appl. Genet.* 128, 1039–1047. doi: 10.1007/s00122-015-2488-y

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B-Methodol.* 57, 289–300.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brown, A. F., Yousef, G. G., Chebrolu, K. K., Byrd, R. W., Everhart, K. W., Thomas, A., et al. (2014). High-density single nucleotide polymorphism (SNP) array mapping in *Brassica oleracea*: identification of QTL associated with carotenoid variation in broccoli florets. *Theor. Appl. Genet.* 127, 2051–2064. doi: 10.1007/s00122-014-2360-5

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., et al. (2009). The genetic architecture of maize flowering time. *Science* 325, 714–718. doi: 10.1126/science.1174276

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Churchill, G. A., and Diverge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138, 963–971.

Dabney, A., and Storey, J. D. (2004). Q-value estimation for false discovery rate control. *Medicine* 344, 539–548.

Dong, Y., and Wang, Y. Z. (2015). Seed shattering: from models to crops. *Front. Plant Sci.* 6:476. doi: 10.3389/fpls.2015.00476

Dong, Y., Yang, X., Liu, J., Wang, B. H., Liu, B. L., and Wang, Y. Z. (2014). Pod dehiscence resistance associated with domestication is mediated by a NAC gene in soybean. *Nat.Commun.* 5, 3352. doi: 10.1038/ncomms4352

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Ferrándiz, C., Liljegren, S. J., and Yanofsky, M. F. (2000). Negative regulation of the SHATTERPROOF genes by FRUITFULL during *Arabidopsis* fruit development. *Science* 289, 436–438. doi: 10.1126/science.289.5478.436

Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374. doi: 10.1146/annurev.arplant.54.031902.134907

Fu, Y., Wei, D. Y., Dong, H. L., He, Y. J., Cui, Y. X., Mei, J. Q., et al. (2015). Comparative quantitative trait loci for silique length and seed weight in *Brassica napus*. *Sci. Rep.* 5:14407. doi: 10.1038/srep14407

Funatsuki, H., Suzuki, M., Hirose, A., Inaba, H., Yamada, T., Hajika, M., et al. (2014). Molecular basis of a shattering resistance boosting global dissemination of soybean. *Proc. Natl. Acad. Sci. U.S.A.* 111, 17797–17802. doi: 10.1073/pnas.1417282111

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424

Gajardo, H., Wittkop, B., Soto-Cerda, B., Higgins, E., Parkin, I. P., Snowdon, R., et al. (2015). Association mapping of seed quality traits in *Brassica napus* L. *using GWAS and candidate QTL approaches. Mol. Breed.* 35, 1–19. doi: 10.1007/s11032-015-0340-3

Gao, M., and Zhu, H. (2013). Fine mapping of a major quantitative trait locus that regulates pod shattering in soybean. *Mol. Breed.* 32, 485–491. doi: 10.1007/s11032-013-9868-2

Girin, T., Stephenson, P., Goldsack, C. M., Kempin, S. A., Perez, A., Pires, N., et al. (2010). Brassicaceae *INDEHISCENT* genes specify valve margin cell fate and repress replum formation. *Plant J.* 63, 329–338. doi: 10.1111/j.1365-313X.2010.04244.x

Heisler, M. G., Atkinson, A., Bylstra, Y. H., Walsh, R., and Smyth, D. R. (2001). *SPATULA*, a gene that controls development of carpel margin tissues in Arabidopsis, encodes a bHLH protein. *Development* 128, 1089–1098.

Hossain, S., Kadkol, G., Raman, R., Salisbury, P., and Raman, H. (2011). "Breeding *Brassica napus* for shatter resistance," in *Plant Breeding,* ed I. Y. Abdurakhmonov. Available online at: http://www.intechopen.com/articles/show/title/breeding-brassica-napus-for-shatter-resistance

Hu, Z. Y., Hua, W., Huang, S. M., Yang, H. L., Zhan, G. M., Wang, X. F., et al. (2012). Discovery of pod shatter-resistant associated SNPs by deep sequencing of a representative library followed by bulk segregant analysis in rapeseed. *PLoS ONE* 7:e34253. doi: 10.1371/journal.pone.0034253

Hua, J. P., Xing, Y. Z., Xu, C. G., Sun, X. L., Yu, S. B., and Zhang, Q. F. (2002). Genetic dissection of an elite rice hybrid revealed that heterozygotes are not always advantageous for performance. *Genetics* 162, 1885–1895.

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233

Jaradat, M. R., Ruegger, M., Bowling, A., Butler, H., and Cutler, A. J. (2014). A comprehensive transcriptome analysis of silique development and dehiscence in Arabidopsis and Brassica integrating genotypic, interspecies and developmental comparisons. *GM Crops Food* 5, 302–320. doi: 10.4161/21645698.2014.947827

Kadkol, G. P., Halloran, G. M., and MacMillan, R. H. (1985). Evaluation of *Brassica* genotypes for resistance to shatter. *II. Variation in siliqua strength within and between accessions. Euphytica* 34, 915–924. doi: 10.1007/BF00035431

Kadkol, G. P., Halloran, G. M., and MacMillan, R. H. (1986). Inheritance of siliqua strength in *Brassica campestris* L. I. Studies of F_2 and backcross populations. *Can. J. Genet. Cytol.* 28, 365–373. doi: 10.1139/g86-054

Kadkol, G. P., MacMillan, R. H., Burrow, R. P., and Halloran, G. M. (1984). Evaluation of *Brassica* genotypes for resistance to shatter. *I. Development of a laboratory test. Euphytica* 33, 63–73. doi: 10.1007/BF00022751

Konishi, S., Izawa, T., Lin, S. Y., Ebana, K., Fukuta, Y., Sasaki, T., et al. (2006). An SNP caused loss of seed shattering during rice domestication. *Science* 312, 1392–1396. doi: 10.1126/science.1126410

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Krill, A. M., Kirst, M., Kochian, L. V., Buckler, E. S., and Hoekenga, O. A. (2010). Association and linkage analysis of aluminum tolerance genes in maize. *PLoS ONE* 5:e9958. doi: 10.1371/journal.pone.0009958

Lewis, M. W., Leslie, M. E., and Liljegren, S. J. (2006). Plant separation: 50 ways to leave your mother. *Curr. Opin. Plant Biol.* 9, 59–65. doi: 10.1016/j.pbi.2005.11.009

Li, C., Zhou, A., and Sang, T. (2006). Rice domestication by reducing shattering. *Science* 311, 1936–1939. doi: 10.1126/science.1123604

Li, N., Shi, J., Wang, X., Liu, G., and Wang, H. Z. (2014). A combined linkage and regional association mapping validation and fine mapping of two major pleiotropic QTLs for seed weight and silique length in rapeseed (*Brassica napus* L.). *BMC Plant Biol.* 14:114. doi: 10.1186/1471-2229-14-114

Liljegren, S. J., Ditta, G. S., Eshed, Y., Savidge, B., Bowman, J., and Yanofsky, M. F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in *Arabidopsis*. *Nature* 404, 766–770. doi: 10.1038/35008089

Liu, J., Hua, W., Hu, Z. Y., Yang, H. L., Zhang, L., Li, R. J., et al. (2015). Natural variation in ARF18 gene simultaneously affects seed weight and silique length in polyploid rapeseed. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5123–E5132. doi: 10.1073/pnas.1502160112

Liu, J., Mei, D. S., Li, Y. C., Cui, J. C., Wang, H., Peng, P. F., et al. (2013). Combining ability and breeding potential of rapeseed elite lines for pod shatter resistance. *J. Integr. Agric.* 12, 552–555. doi: 10.1016/s2095-3119(13)60256-7

Liu, L., Qu, C., Wittkop, B., Yi, B., Xiao, Y., He, Y. et al. (2013). A High-Density SNP Map for Accurate Mapping of Seed Fibre QTL in *Brassica napus* L. *PLoS ONE* 8:e83052. doi: 10.1371/journal.pone.0083052

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A. P., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930. doi: 10.1038/ncomms4930

Liu, Z. S. (1994). Genetic and breeding studies on distant hybridization in rapeseed. II. Cross compatibility between *Brassica napus* and *B. juncea* and their F1s. *Oil Crops China* 16, 1–5.

Loiselle, B. A., Sork, V. L., Nason, J., and Graham, C. (1995). Spatial genetic-structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am. J. Bot.* 82, 1420–1425. doi: 10.2307/2445869

Mauricio, R. (2001). Mapping quantitative trait loci in plants: uses and caveats for evolutionary biology. *Nat. Rev. Genet.* 2, 370–381. doi: 10.1038/35072085

Mongkolporn, O., Kadkol, G. P., Pang, E. C. K., and Taylor, P. W. J. (2003). Identification of RAPD markers linked to recessive genes conferring siliqua shatter resistance in *Brassica rapa*. *Plant Breed.* 122, 479–484. doi: 10.1046/j.0179-9541.2003.00910.x

Morgan, C. L., Bruce, D. M., Child, R., Ladbrooke, Z. L., and Arthur, A. E. (1998). Genetic variation for pod shatter resistance among lines of oilseed rape developed from synthetic B-Napus. *Field Crops Res.* 58. 153–165. doi: 10.1016/S0378-4290(98)00099-9

Morgan, C. L., Ladbrooke, Z. L., Bruce, D. M., Child, R., and Arthur, A., E. (2000). Breeding oilseed rape for pod shattering resistance. *J. Agric. Sci.* 135, 347–359. doi: 10.1017/s0021859699008424

Okushima, Y., Overvoorde, P. J., Arima, K., Alonso, J. M., Chan, A., Chang, C., et al. (2005). Functional genomic analysis of the *AUXIN RESPONSE FACTOR* gene family members in *Arabidopsis thaliana*: unique and overlapping functions of *ARF7* and *ARF19*. *Plant Cell* 17, 444–463. doi: 10.1105/tpc.104.028316

Østergaard, L. (2009). Don't 'leaf' now. The making of a fruit. *Curr. Opin. Plant Biol.* 12, 36–41. doi: 10.1016/j.pbi.2008.09.011

Parkin, I. A. P., Koh, C., Tang, H., Robinson, S., Kagale, S., Clarke, W., et al. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome Biol.* 15:R77. doi: 10.1186/gb-2014-15-6-r77

Peng, P., Li, Y. C., Mei, D., Liu, D., Fu, L., Wang, H., et al. (2013). Optimization and experiment of assessment method for pod shatter resistance in *Brassica napus* L. *Trans. Chin. Soc. Agric. Eng.* 29, 19–25. doi: 10.3969/j.issn.1002-6819.2013.21.003

Price, J. S., Hobson, R. N., Neale, M. A., and Bruce, D. M. (1996). Seed losses in commercial harvesting of oilseed rape. *J. Agric. Eng. Res.* 65, 183–191. doi: 10.1006/jaer.1996.0091

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Pu, H. M., Long, W. H., Gao, J. Q., and Hu, M. L. (2013). Silique shatter resistance and correlation analysis in *Brassica napus*. *Oil Crops China* 35, 469–475. doi: 10.7505/j.issn.1007-9084.2013.05.002

Qian, W., Chen, X., Fu, D., Zou, J., and Meng, J. (2005). Intersubgenomic heterosis in seed yield potential observed in a new type of *Brassica napus* introgressed with partial *Brassica rapa* genome. *Theor. Appl. Genet.* 110, 1187–1194. doi: 10.1007/s00122-005-1932-9

Rajani, S., and Sundaresan, V. (2001). The Arabidopsis myc/bHLH gene *ALCATRAZ* enables cell separation in fruit dehiscence. *Curr. Biol.* 11, 1914–1922. doi: 10.1016/S0960-9822(01)00593-0

Raman, H., Raman, R., Coombes, N., Song, J., Prangnell, R., Bandaranayake, C., et al. (2016). Genome-wide association analyses reveal complex genetic architecture underlying natural variation for flowering time in canola. *Plant Cell Environ.* 39, 1228–1239. doi: 10.1111/pce.12644

Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., et al. (2014). Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLoS ONE* 9:e101673. doi: 10.1371/journal.pone.0101673

Saghai-Maroof, M. A., Soliman, K. M., Jorgensen, R. A., and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population-dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 81, 8014–8018. doi: 10.1073/pnas.81.24.8014

SAS Institute Inc. (1999). *SAS Online Doc@, version 8.1*. Cary, NC.

Schiessl, S., Samans, B., Hüttel, B., Reinhardt, R., and Snowdon, R. J. (2014). Capturing sequence variation among flowering-time regulatory gene homologues in the allopolyploid crop species *Brassica napus*. *Front. Plant Sci.* 5:404. doi: 10.3389/fpls.2014.00404

Sorefan, K., Girin, T., Liljegren, S. J., Ljung, K., Robles, P., Galvan-Ampudia, C. S., et al. (2009). A regulated auxin minimum is required for seed dispersal in *Arabidopsis*. *Nature* 459, 583–586. doi: 10.1038/nature07875

Stam, P. (1993). Construction of integrated genetic linkage maps by means of a new computer package: joinmap. *Plant J.* 3, 739–744. doi: 10.1111/j.1365-313X.1993.00739.x

Sun, G., Zhu, C. M., Kramer, H. S., Yang, S., Song, W., Piepho, H. P. et al. (2010). Variation explained in mixed-model association mapping. *Heredity* 105, 333–340. doi: 10.1038/hdy.2010.11

Trick, M., Long, Y., Meng, J., and Bancroft, I. (2009). Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnol. J.* 7, 334–346. doi: 10.1111/j.1467-7652.2008.00396.x

Wang, J. K., Li, H. H., Zhang, L. Y., Li, C. H., and Meng, L. (2014). *Users' Manual of QTL IciMapping V4.0*. Beijing. Available online at: http://www.isbreeding.net/software/?type=detail&id=14

Wang, S., Basten, C. J., and Zeng, Z. B. (2007). *Windows QTL Cartographer 2.5*. Raleigh, NC: Department of Statistics, North Carolina State University. Available online at: http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S. Y., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Wang, X., Yu, K., Li, H., Peng, Q., Chen, F., Zhang, W., et al. (2015). High-density SNP map construction and QTL identification for the apetalous character in *Brassica napus* L. *Front. Plant Sci.* 6:1164. doi: 10.3389/fpls.2015.01164

Wei, W. H., Li, Y. C., Wang, L. J., Liu, S. Y., Yan, X. H., Mei, D. S., et al. (2010). Development of a novel *Sinapis arvensis* disomic addition line in *Brassica napus* containing the restorer gene for Nsa CMS and improved resistance to *Sclerotinia sclerotiorum* and pod shattering. *Theor. Appl. Genet.* 120, 1089–1097. doi: 10.1007/s00122-009-1236-6

Wen, Y. C., Fu, T., Tu, J., Ma, C., Shen, J., Zhang, S., et al. (2008). Screening and analysis of resistance to silique shattering in rape (*Brassica napus* L.). *Acta Agron. Sin.* 34, 163–166. doi: 10.3724/sp.j.1006.2008.00163

Wen, Y. C., Zhang, S. F., Yi, B., Wen, J., Wang, J. P., et al. (2013). Identification of QTLs involved in pod-shatter resistance in *Brassica napus* L. *Crop Pasture Sci.* 63, 1082–1089. doi: 10.1071/CP12318

Xu, J., Qian, X., Wang, X., Li, R., Cheng, X., Yang, Y., et al. (2010). Construction of an integrated genetic linkage map for the A genome of *Brassica napus* using SSR markers derived from sequenced BACs in *B. rapa*. *BMC Genomics* 11:594. doi: 10.1186/1471-2164-11-594

Yoon, J., Cho, L. H., Kim, S. L., Choi, H., Koh, H. J., and An, G. (2014). The BEL1- type homeobox gene SH5 induces seed shattering by enhancing abscission- zone development and inhibiting lignin biosynthesis. *Plant J.* 79, 717–728. doi: 10.1111/tpj.12581

Yu, J. M., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

Zhou, Y., Lu, D., Li, C., Luo, J., Zhu, B. F., Zhu, J., et al. (2012). Genetic control of seed shattering in rice by the APETALA2 transcription factor shattering abortion1. *Plant Cell* 24, 1034–1048. doi: 10.1105/tpc.111.094383

Zou, J., Zhu, J., Huang, S., Tian, E., Xiao, Y., Fu, D., et al. (2010). Broadening the avenue of intersubgenomic heterosis in oilseed *Brassica*. *Theor. Appl. Genet.* 120, 283–290. doi: 10.1007/s00122-009-1201-4

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Liu, Wang, Wang, Wang, Zhou, Mei, Cheng, Yang, Raman and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 July 2016
doi: 10.3389/fpls.2016.01054

Identification of microRNAs and Their Target Genes Explores miRNA-Mediated Regulatory Network of Cytoplasmic Male Sterility Occurrence during Anther Development in Radish (*Raphanus sativus* L.)

Wei Zhang[1†], *Yang Xie*[1†], *Liang Xu*[1], *Yan Wang*[1], *Xianwen Zhu*[2], *Ronghua Wang*[1], *Yang Zhang*[1], *Everlyne M. Muleke*[1] *and Liwang Liu*[1*]

[1] National Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing, China, [2] Department of Plant Sciences, North Dakota State University, Fargo, ND, USA

MicroRNAs (miRNAs) are a type of endogenous non-coding small RNAs that play critical roles in plant growth and developmental processes. Cytoplasmic male sterility (CMS) is typically a maternally inherited trait and widely used in plant heterosis utilization. However, the miRNA-mediated regulatory network of CMS occurrence during anther development remains largely unknown in radish. In this study, a comparative small RNAome sequencing was conducted in floral buds of CMS line 'WA' and its maintainer line 'WB' by high-throughput sequencing. A total of 162 known miRNAs belonging to 25 conserved and 24 non-conserved miRNA families were isolated and 27 potential novel miRNA families were identified for the first time in floral buds of radish. Of these miRNAs, 28 known and 14 potential novel miRNAs were differentially expressed during anther development. Several target genes for CMS occurrence-related miRNAs encode important transcription factors and functional proteins, which might be involved in multiple biological processes including auxin signaling pathways, signal transduction, miRNA target silencing, floral organ development, and organellar gene expression. Moreover, the expression patterns of several CMS occurrence-related miRNAs and their targets during three stages of anther development were validated by qRT-PCR. In addition, a potential miRNA-mediated regulatory network of CMS occurrence during anther development was firstly proposed in radish. These findings could contribute new insights into complex miRNA-mediated genetic regulatory network of CMS occurrence and advance our understanding of the roles of miRNAs during CMS occurrence and microspore formation in radish and other crops.

Keywords: radish (*Raphanus sativus* L.), cytoplasmic male sterility, microRNA, target gene, qRT-PCR, high-throughput sequencing

OPEN ACCESS

Edited by:
Narendra Tuteja,
International Centre for Genetic Engineering and Biotechnology, India

Reviewed by:
Lijun Chai,
Huazhong Agricultural University, China
Maoteng Li,
Huazhong University of Science and Technology, China
Anca Macovei,
University of Pavia, Italy

***Correspondence:**
Liwang Liu
nauliulw@njau.edu.cn

[†]These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *31 December 2015*
Accepted: *05 July 2016*
Published: *22 July 2016*

Citation:
Zhang W, Xie Y, Xu L, Wang Y, Zhu X, Wang R, Zhang Y, Muleke EM and Liu L (2016) Identification of microRNAs and Their Target Genes Explores miRNA-Mediated Regulatory Network of Cytoplasmic Male Sterility Occurrence during Anther Development in Radish (Raphanus sativus L.). Front. Plant Sci. 7:1054. doi: 10.3389/fpls.2016.01054

INTRODUCTION

MicroRNAs (miRNAs) are a type of endogenous non-coding small RNAs of ∼21–24 nucleotides that are known to be important negative regulators of gene expression at transcriptional and post-transcriptional level by mediating mRNA degradation or translational repression (Voinnet, 2009). In plants, primary miRNAs (pri-miRNAs) are transcribed from nuclear-encoded *MIR* genes by RNA polymerase II and cleaved by Dicer-like1 (DCL1) assisted by the dsRNA binding protein HYL1 to generate miRNA:miRNA* duplexes called pre-miRNAs (Jones-Rhoades et al., 2006; Kurihara et al., 2006; Ruiz-Ferrer and Voinnet, 2009). The duplexes are then methylated by HEN1 and one of the strands combines with the argonaute protein1 (AGO1) to form the RNA-induced silencing complex (RISC), which regulates gene expression through mRNA degradation with nearly perfect complementarity or translational repression with partial complementarity (Yu et al., 2005; Jones-Rhoades et al., 2006; Bodersen et al., 2008).

Cytoplasmic male sterility (CMS) is a maternally inherited trait in plant, which is unable to produce functional pollen, and is a widely observed phenomenon in nearly 200 species (Brown et al., 2003; Hu et al., 2012). CMS lines have been widely used for the production of F_1 hybrid seeds and utilization of heterosis in many crops, such as cotton, maize, sorghum, wheat, rice, beet, and rapeseed (Schnable and Wise, 1998; Bentolila et al., 2002; Kubo et al., 2011). In addition to its crucial breeding tools, CMS lines also provide important materials for studying anther and pollen development, and cytoplasmic-nuclear interactions (Chen and Liu, 2014). CMS is usually due to the effect of sterilizing factors found in the mitochondrial genome (Touzet and Meyer, 2014). In most cases, CMS can be restored by nuclear-encoded fertility restorer (*Rf*) gene(s), which relies on *Rf* suppressing cytoplasmic dysfunction caused by mitochondrial genes (Eckardt, 2006). High-throughput sequencing now is widely used and has been proven to be an excellent application for the identification of plant miRNAs. As a class of negative regulators, miRNAs have also been identified and characterized during anther development in several plant species, including *Arabidopsis* (Chambers and Shuai, 2009), *Oryza sativa* (Wei et al., 2011; Yan et al., 2015), *Gossypium hirsutum* (Wei et al., 2013), *Brassica juncea* (Yang et al., 2013), and *B. rapa* (Jiang et al., 2014). In *G. hirsutum*, 16 conserved miRNA families were identified during anther development between the Genetic male sterility (GMS) mutant and its wild type. In *O. sativa*, Wei et al. (2011) identified 292 known miRNAs and 75 novel miRNAs from sporophytic tissues and pollen at three developmental stages. Additionally, many CMS occurrence-associated miRNAs have also been identified in some vegetable crops. In *B. rapa*, 54 conserved and eight novel miRNA families involved in pollen development were identified (Jiang et al., 2014). In *B. juncea*, 197 known and 93 new candidate miRNAs during pollen development between CMS line and its maintainer line were also identified (Yang et al., 2013). Although a large number of miRNAs during anther development have been isolated and identified in many crop species, the miRNA-mediated regulatory network of CMS occurrence during anther development remain to be clarified in root vegetable crops.

Radish (*Raphanus sativus* L. $2n = 2x = 18$) is an important annual or biennial root vegetable crop of Brassicaceae family. In recent years, some conserved and novel miRNAs associated with taproot thickening, embryogenesis, flowering-time, and heavy metal stresses had been widely identified in radish (Xu et al., 2013a; Zhai et al., 2014; Nie et al., 2015; Wang et al., 2015; Yu et al., 2015). However, there is little information about the CMS occurrence at the post-transcriptional level in radish. To systematically explore the roles of miRNAs and their targets involved in CMS occurrence during anther development in radish, two small RNA libraries from 'WA' (male sterile line), and 'WB' (maintainer, fertile line) floral buds of radish were constructed. The aims of this study were to identify known and potential novel miRNAs from the two libraries and investigate the dynamic expression patterns of the CMS occurrence-related miRNAs and their targets during anther development in radish plant. Furthermore, the miRNA-mediated regulatory network of CMS occurrence during anther development was constructed in radish. These results would lay a valuable foundation for elucidating the regulatory roles of CMS occurrence-related miRNAs in radish and facilitate further dissection of the molecular mechanisms underlying microspore formation and CMS occurrence in other crops.

MATERIALS AND METHODS

Plant Materials

The radish cytoplasmic male sterile line 'WA' and its maintainer line 'WB' were used as materials in this study. The 'WB' was advanced inbred line through multiple self-pollination for more than 10 generations, while CMS line 'WA' was developed through continuously backcrossing with 'WB' for more than 10 generations. 'WA' had completely aborted anthers without pollen, whereas 'WB' had normal anthers with fertile pollen (**Figure S1**). The materials were planted under normal conditions at Jiangpu Breeding Station of Nanjing Agricultural University, China. According to the cytological characterization of the developmental stages identified with paraffin section technique, the longitudinal length of floral buds reaching 1–1.5, 2–2.5, and 4–5 mm corresponds to the stage of meiosis, tetrad, and early microscope, respectively (**Figure S1**), which was in highly accordance with results of previous studies (Sun et al., 2012, 2013). Floral buds at three stages were independently collected from the two lines with three biological replicates. Each sample was collected from three randomly selected individual plants and immediately frozen in liquid nitrogen and stored at −80°C for further use.

High-Throughput Sequencing of Small RNAs

Total RNAs were extracted from three stages of floral buds of 'WA' and 'WB' using Trizol® Reagent (Invitrogen, USA) according to the manufacturer's protocols, respectively. RNAs from the three different stages were equally pooled and used for

two small RNA libraries (WA and WB) construction according to previously described procedures (Hafner et al., 2008; Xu et al., 2013b). In brief, small RNA fractions of 18–30 nt were separated and purified from total RNA by 15% denaturing polyacrylamide gel electrophoresis. Then the isolated sRNAs were ligated to 5′ and 3′ adaptors and reverse transcribed to cDNA through SuperScript II Reverse Transcriptase (Invitrogen) and amplified by PCR. Finally, sRNA libraries were sequenced by the Solexa sequencer (Illumina) HiSeq^{TM} 2500.

The clean reads were obtained after removing low quality reads, reads with 5′ primer contaminants or poly-A tails, trimming reads smaller than 18 nt or longer than 30 nt. The remaining unique sequences were then mapped to the radish reference sequences including genomic survey sequences (GSS), expressed sequence tag (EST) sequences and the radish mRNA transcriptome sequences (accession number: SRX1671013) using the SOAP2 program (Li et al., 2009; Xu et al., 2013a). Only perfect matched sequences with no more than two mismatches were retained for proceeding analysis. After using BLAST in GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and Rfam 12.0 (http://rfam.xfam.org/) database, the clean reads compared with the non-coding RNAs (rRNAs, tRNAs, snRNAs, and snoRNA) were removed for further analysis. The remaining matched reads were aligned with known miRNAs in miRBase 21 (http://www.mirbase.org/index.shtml) for radish known miRNAs identification. Then, the unannotated reads were used to predict potential novel miRNAs using Mireap software (https://sourceforge.net/projects/mireap/) according to the previous criteria (Meyers et al., 2008). The stem-loop structure of miRNA precursors were folded by Mfold (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (Zuker, 2003).

Differential Expression Analysis of miRNAs between CMS Line and Its Maintainer Line

The frequency of miRNAs from two libraries was normalized to one million by total number of miRNAs per sample (Gao et al., 2012). If the normalized read of a given miRNA is zero, the expression value was set to 0.01 for further use. The differential expression of miRNAs between the two libraries was calculated as: Fold-change $= \log_2$ (WA/WB). The *P*-value was calculated following the previously reported methods (Li et al., 2009; Zhai et al., 2014). The miRNAs with $P \leq 0.05$ and fold-change ≥ 1 or ≤ -1 were considered as up- or down-regulated miRNAs between the two libraries during anther development, respectively.

Prediction and Annotation of Potential Targets for CMS Occurrence-Related miRNAs

The potential target genes of the identified miRNAs were predicted by the plant small RNA target analysis server (psRNATarget; http://plantgrn.noble.org/psRNATarget/) (Dai and Zhao, 2011). The criteria used for target prediction in plants were performed following previous methods (Allen et al., 2005). To understand the biological functions of the targets, gene ontology (GO) analysis were performed by Blast2GO program on the basis of the BLAST searching against the available Nr database in NCBI. In addition, KEGG Orthology Based Annotation System (KOBAS2.0; http://kobas.cbi.pku.edu.cn/home.do) was applied to predict the biological functions of target genes (Xie et al., 2011). Based on the differentially expressed miRNAs and their corresponding targets, the miRNA-targets regulatory network was constructed using Cytoscape_v3.2.1 software (Smoot et al., 2011).

qRT–PCR Validation

Quantitative reverse transcription-PCR (qRT–PCR) was employed to evaluate the validity of small RNA sequencing and also to analyze the expression patterns of miRNAs and their targets during different stages. miRNAs and total RNAs were extracted from samples and reverse-transcribed to cDNA using the One Step Primer Script® miRNA cDNA Synthesis Kit (Takara Bio Inc., Dalian, China) and SuperScript® III Reverse Transcriptase (Invitrogen, USA) following the manufacturer's instructions, respectively. All reactions were performed on a BioRad iQ5 sequence detection system (BIO-RAD) and carried out in a total volume of 20 μl including 0.2 μM primer pairs, 2 μl diluted cDNA, and 10 μl 2 × SYBR Green PCR Master Mix (TaKaRa). The PCR amplification reaction was performed following the previous reports (Zhai et al., 2014). The 5.8S ribosomal RNA (rRNA) was used as the reference gene for normalization. All reactions were done in triplicate, the $2^{-\Delta\Delta C_T}$ method was used to calculate the relative expression data (Livak and Schmittgen, 2001). The statistical analysis was performed using SPSS 20 software (SPSS Inc., USA) with Duncan's multiple range test at the 5% level of significance. The primers for qRT–PCR were showed in **Table S1**.

RESULTS

Overview Analysis of Sequences from Small RNA Libraries

To identify known and potential novel miRNAs involved in anther development and CMS occurrence, we constructed two small RNA libraries from the floral buds of 'WA' and 'WB' line. A total of 43,068,458 raw reads were obtained from the two sRNA libraries. After filtering low quality reads, adapter contaminants, and reads smaller than 18 nucleotides, we obtained 20,287,225 (representing 5,528,061 unique sequences), and 21,989,236 (representing 5,682,107 unique sequences) clean reads from WA and WB library, respectively (**Table S2**). Of these reads, 13.84% were WA library-specific with 42.68% unique sRNAs, 13.88% were WB library-specific with 44.24% unique sRNAs, and 72.28% were present in both with 13.08% unique sRNAs (**Table S3**).

By comparing with the NCBI GenBank and Rfam databases, these clean reads that matched non-coding sRNAs including rRNAs, snoRNAs, snRNAs, and tRNAs were eliminated. After that, 27,092 (WA) and 27,764 (WB) unique sequences were acquired by querying the unique reads against miRBase 21 (**Table 1**). The remaining 5,388,388 (WA) and 5,511,728 (WB) unannotated unique reads were used for identification of potential novel miRNAs (**Table 1**). The length distribution of

TABLE 1 | Distribution of small RNAs among different categories in radish.

Category	WA		WB	
	Unique sRNAs	Total sRNAs	Unique sRNAs	Total sRNAs
Total	5528061 (100%)	20287225 (100%)	5682107 (100%)	21989236 (100%)
miRNA	27092 (0.49%)	1453994 (7.17%)	27764 (0.49%)	1521955 (6.92%)
rRNA	92089 (1.67%)	2685339 (13.24%)	117885 (2.07%)	3446274 (15.67%)
snRNA	6099 (0.11%)	22072 (0.11%)	6602 (0.12%)	27969 (0.13%)
snoRNA	3383 (0.06%)	10011 (0.05%)	4120 (0.07%)	11837 (0.05%)
tRNA	11010 (0.20%)	1157890 (5.71%)	14008 (0.25%)	514526 (2.34%)
unannotated	5388388 (97.47%)	14957919 (73.73%)	5511728 (97%)	16466675 (74.89%)

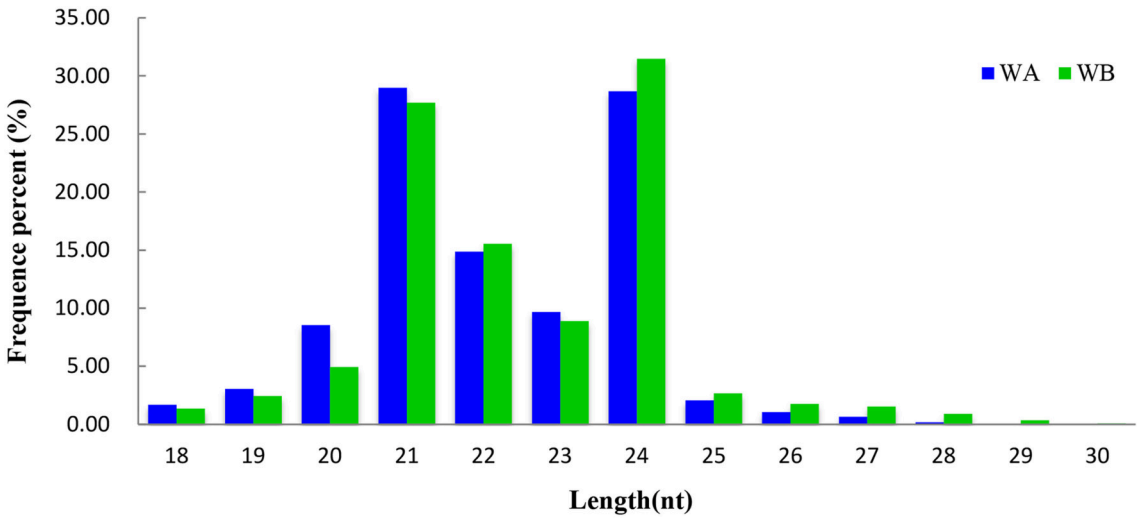


FIGURE 1 | Length distribution of small RNAs in WA and WB library.

sRNA reads ranged from 18 to 30 nt in both libraries (**Figure 1**), and the most abundant sequences in the two libraries ranged from 20 to 24 nt, which is the representative size range of products cleaved by DCLs (Henderson et al., 2006). The most abundant sRNAs in WA and WB library was 21 and 24 nt long, which accounted for 28.97 and 31.47%, respectively.

Identification of Known miRNAs during Anther Development

To identify known miRNAs from the two libraries, the unique sRNA reads were aligned to known miRNA precursors, and mature miRNA sequences in miRBase 21, allowing a maximum of two mismatches. A total of 124 unique reads belonging to 25 conserved miRNA families were identified in the two libraries (**Table 2**). The distribution of conserved miRNA family members was analyzed (**Figure S2**). A large part of conserved miRNA families had members of more than three, and miR165/166 family possessed the largest member of 17, followed by miR156/157, and miR169 with 14 and 11 members, respectively. However, some conserved miRNA families including miR158, miR161, miR391, miR395, miR397 miR398, and miR403 had only one or two members. In addition, 38 unique reads belonging to 24 non-conserved miRNA families were also discovered in these two libraries, which contained fewer members as compared with conserved miRNAs (**Figure S2**).

The number of miRNA reads differed greatly in the two libraries (**Figure S3**). For instance, miR156/157 presented the highest expression abundance with 410,237 in WA library, while miRNA165/166 displayed the highest expression of 405,255 copies in WB library. Several miRNA families such as miR167, miR168, miR2118, and miR2199 also displayed extraordinarily high abundance in both libraries, while some other miRNA families (miR400, miR828, miR829, miR831, and miR858) were expressed with relatively low levels of expression with no more than 100 reads in WA and WB library. In addition, the expression levels of different members of the same miRNA family varied drastically (**Table S4**).

Identification of Potential Novel miRNAs in Floral Buds

A total of 30 precursor sequences and 27 novel miRNA families were identified in the two libraries (**Table S5**). The secondary

TABLE 2 | Known miRNA families and their expression abundance in WA and WB library.

Family	Number of members	miRNA reads		Total reads	Ratio (WA/WB)
		WA	WB		
CONSERVED miRNA					
miR156/157	14	410237	304695	714932	1.35
miR158	2	15565	7641	23206	2.04
miR159	4	1726	1157	2883	1.49
miR160	6	13463	18893	32356	0.71
miR161	1	152	0	152	–
miR162	3	1211	1379	2590	0.88
miR164	5	9831	9100	18931	1.08
miR165/166	17	321161	405255	726416	0.79
miR167	8	318811	276963	595774	1.15
miR168	3	137677	135624	273301	1.02
miR169	11	36871	37179	74050	0.99
miR171	6	3241	4824	8065	0.67
miR172	7	9866	13246	23112	0.74
miR319	4	508	976	1484	0.52
miR390	5	7594	10560	18154	0.72
miR391	2	3769	3177	6946	1.19
miR393	3	1789	1643	3432	1.09
miR394	3	358	348	706	1.03
miR395	2	45236	44	45280	1028.09
miR396	6	3751	2853	6604	1.31
miR397	1	13	60	73	0.22
miR398	2	223	285	508	0.78
miR399	3	108	80	188	1.35
miR403	2	1455	1306	2761	1.11
miR408	4	3431	7780	11211	0.44
NON-CONSERVED miRNA					
miR400	2	43	33	76	1.30
miR447	1	189	245	434	0.77
miR482	2	2073	35	2108	59.23
miR529	3	68	55	123	1.24
miR535	2	2229	83	2312	26.86
miR824	2	651	681	1332	0.96
miR825	1	354	479	833	0.74
miR827	3	1932	2558	4490	0.76
miR828	1	5	71	76	0.07
miR829	1	36	98	134	0.37
miR831	1	87	58	145	1.50
miR845	3	5094	5262	10356	0.97
miR854	1	640	468	1108	1.37
miR858	1	0	26	26	0.00
miR859	1	671	0	671	–
miR860	1	248	215	463	1.15
miR948	1	0	336	336	0.00
miR1878	1	0	1584	1584	0.00
miR1885	2	1961	4902	6863	0.40
miR2111	4	91	206	297	0.44
miR2118	1	42998	60406	103404	0.71
miR2199	1	42085	44471	86556	0.95
miR3444	1	0	452	452	0.00
miR5654	1	1298	1705	3003	0.76

structures of these predicted novel miRNA precursors were displayed in **Figure S4**. In addition to secondary structure prediction, identification of complementary sequences of the mature miRNAs is another way to provide forceful evidences for these predicted novel miRNAs (Meyers et al., 2008). Out of these potential novel miRNAs, only seven miRNAs with mature and complementary miRNA*s were detected as the novel miRNA candidates (**Table 3**). In the present study, the length of these mature miRNAs ranged from 20 to 23 nt, with a distribution peak at 21 nt (60.0%). Furthermore, the length of these potential novel miRNA precursors ranged from 72 to 255 nt with the average length of 142.6 nt. The minimum free energy (MFE) value ranged from −97.83 to −22.9 kcal/mol with an average value of −48.32 kcal/mol. In addition, nine potential novel miRNAs were expressed in both libraries, while a total of 14 and 8 potential novel miRNAs were WA library-specific, and WB library-specific, respectively (**Table S5**). Most of these potential novel miRNAs had relatively low expression levels when compared with known miRNAs, and the expression levels of miRNA* sequences were obviously less than those of their corresponding mature miRNAs, which was consistent with the viewpoint that miRNA* strands degraded quickly during the biogenesis of mature miRNAs (Rajagopalan et al., 2006).

Identification of CMS Occurrence-Related miRNAs during Anther Development in Radish

To identify miRNAs involved in CMS occurrence during anther development in radish, the differential expression of miRNAs in WA, and WB library was analyzed. Based on these rigorous set of criteria above, a total of 28 known and 14 potential novel miRNAs were differentially expressed during anther development (**Figure 2**, **Table S6**). Among them, 17 miRNAs including 11 known and 6 novel ones were up-regulated, and 25 miRNAs including 17 known and 8 novel ones were down-regulated. Of these, 15 miRNAs were differentially expressed at a ratio greater than 10-fold, including 13 known, and two novel miRNAs. Especially, two miRNAs, miR395x (17.77-fold) and rsa-miRn3 (11.09-fold) were the most significantly up-regulated known and novel miRNA, respectively (**Figure 2**). In addition, many of these CMS occurrence-related miRNAs including miR169m, miR171b-3p, miR396b, miR482c-5p, miR1878-3p, and miR3444a-5p were confined to be expressed only in the WA library, whereas miR171a-3p, miR396a, miR482a-5p, and miR859 were only detected in the WB library. The findings suggested that these miRNAs may play critical roles during anther development in radish.

Target Prediction of CMS Occurrence-Related miRNAs in Radish

Target prediction is a prerequisite to understand the biological functions of miRNAs during anther development. In this study, a total of 489 target transcripts were predicted for all the identified miRNAs in radish (**Tables S7**, **S8**). To further understand the biological functions of miRNAs, the annotation of these target transcripts were classified into three GO ontologies using the Blast2GO program (http://www.blast2go.com), including 21 biological processes, 12 cellular components, and 10 molecular functions (**Figure 3**). The main terms in biological processes were "cellular process" (GO: 0009987), "metabolic process" (GO: 0008152), "single-organism process" (GO: 0044699), and "biological regulation" (GO: 0065007). In regard to cellular components, "cell" (GO: 0005623), "cell part" (GO: 0044464), and "organelle" (GO: 0043226) were the three most abundant terms. In addition, "binding" (GO: 0005488) and "catalytic

TABLE 3 | Novel miRNAs with their complementary miRNA*s during anther development in radish.

miRNA name	Reads		Mature sequence (5′–3′)	Arm	Size	LP (nt)	MFE (kcal/mol)	miRNA location
	WA	WB						
rsa-miRn3	443	0	TATTCCGACGACAATTCCGACG	5′	22	100	–56.31	CL2831.Contig4,Contig6
rsa-miRn3*	8	0	TCGGAATTCCGTCGGAATATA	3′	21	100	–56.31	CL2831.Contig4,Contig6
rsa-miRn4	519	637	AATGTATGTAGTCCAATCTAT	5′	21	117	–66	CL2870.Contig2
rsa-miRn4*	18	6	ACATTGGACTACATATATTAC	3′	21	117	–66	CL2870.Contig2
rsa-miRn5	1141	1452	GCTTCCATATCTAGCAGTAGG	5′	21	184	–75.8	CL2916.Contig2
rsa-miRn5*	6	12	TACCGATAGATGTGGAAGCGT	3′	21	184	–75.8	CL2916.Contig2
rsa-miRn7	21141	24827	TTTGCGTGAGTATGTGGATGT	5′	21	119	–49	CL4600.Contig2
rsa-miRn7*	32	42	ATCCACATACTCACGAAAATC	3′	21	119	–49	CL4600.Contig2
rsa-miRn9a	771	1062	CGTTCAGTTCTCCTTTTGCTTC	5′	22	106	–47.24	Rsa#S43006900
rsa-miRn9a*	7	28	AGCAAACGAGAATTGAACGGA	3′	21	106	–47.24	Rsa#S43006900
rsa-miRn19	0	186	GAACGATATAAAAGATCATGGA	5′	22	105	–30.2	CL6156.Contig1, Contig2
rsa-miRn19*	0	44	TATGGCCTTTATATCGTATTCG	3′	22	105	–30.2	CL6156.Contig1, Contig2
rsa-miRn24	0	25	GGTGCAGTTCGGGACTGATTG	5′	21	110	–48.8	FD955742
rsa-miRn24*	0	10	ATTGGCTCCCGCCTTGCATCAA	3′	22	110	–48.8	FD955742

LP (nt), The length of precursor; MFE (kcal/mol), Minimum free energy.

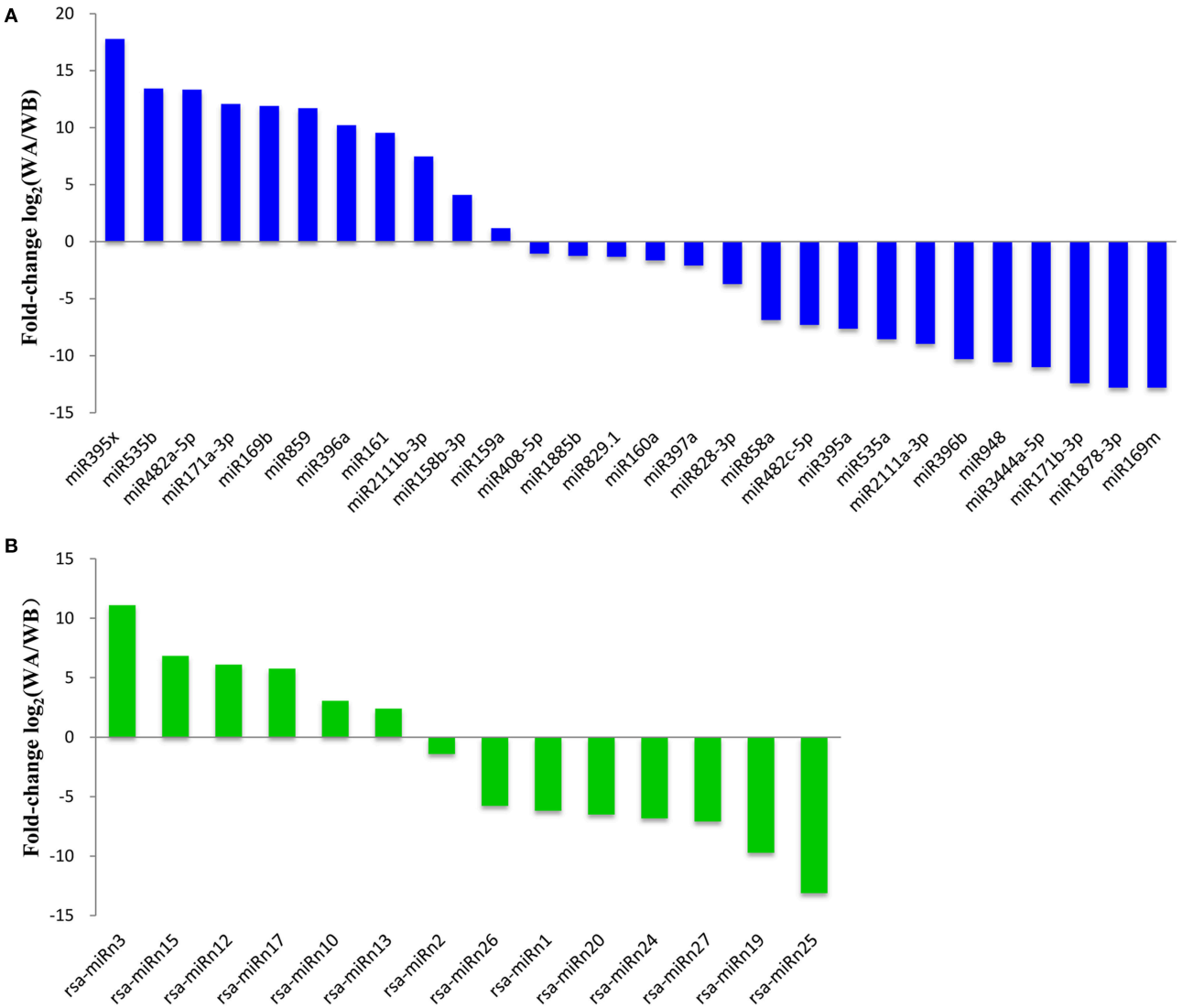


FIGURE 2 | Comparative relative expression of differentially expressed known (A) and potential novel (B) miRNAs between WA and WB library from radish floral buds.

activity" (GO: 0003824) were the most abundant subcategories in the molecular functions.

To understand the biological functions of the isolated miRNAs in radish, the miRNA-cleaved mRNAs during anther development were identified. In this study, 489 potential target sequences for 53 known, 16 potential novel and 84 unclassified non-conserved miRNAs from the transcripts of WA and WB library were further annotated by BLAST search against Arabidopsis sequences using KOBAS 2.0 program (**Tables S7**, **S8**). Among these predicted targets, a large proportion of them are known transcription factor families such as auxin response factors (*ARFs*), basic-leucine zippers (*bZIPs*), myb domain proteins (*MYBs*), and squamosa promoter-binding proteins (*SPLs*), which could play essential roles in anther development and CMS occurrence of radish. Moreover, several target genes encoding functional proteins play roles in a broad range of biological processes including agamous-like MADS-box protein 16 (*AGL16*), argonaute (*AGO*), F-box protein (*F-box*), NAC domain containing protein 96 (*NAC096*), pentatricopeptide repeat-containing protein (*PPR*), and protein TRANSPORT INHIBITOR RESPONSE 1 (*TIR1*) (**Tables 4**, **S7**). To gain further insight into the correlations between miRNAs and their targets, the miRNA-targets regulatory network was constructed (**Figure S5**, **Table S9**). Among them, 26 miRNAs including 19 known and 7 potential novel ones, and 87 unique targets formed a total of 93 miRNA–targets pairs with negatively correlated expression during anther development. In general, these results suggested that the differentially expressed miRNAs may play fundamental regulatory roles in diverse aspects of biological processes during anther development of radish.

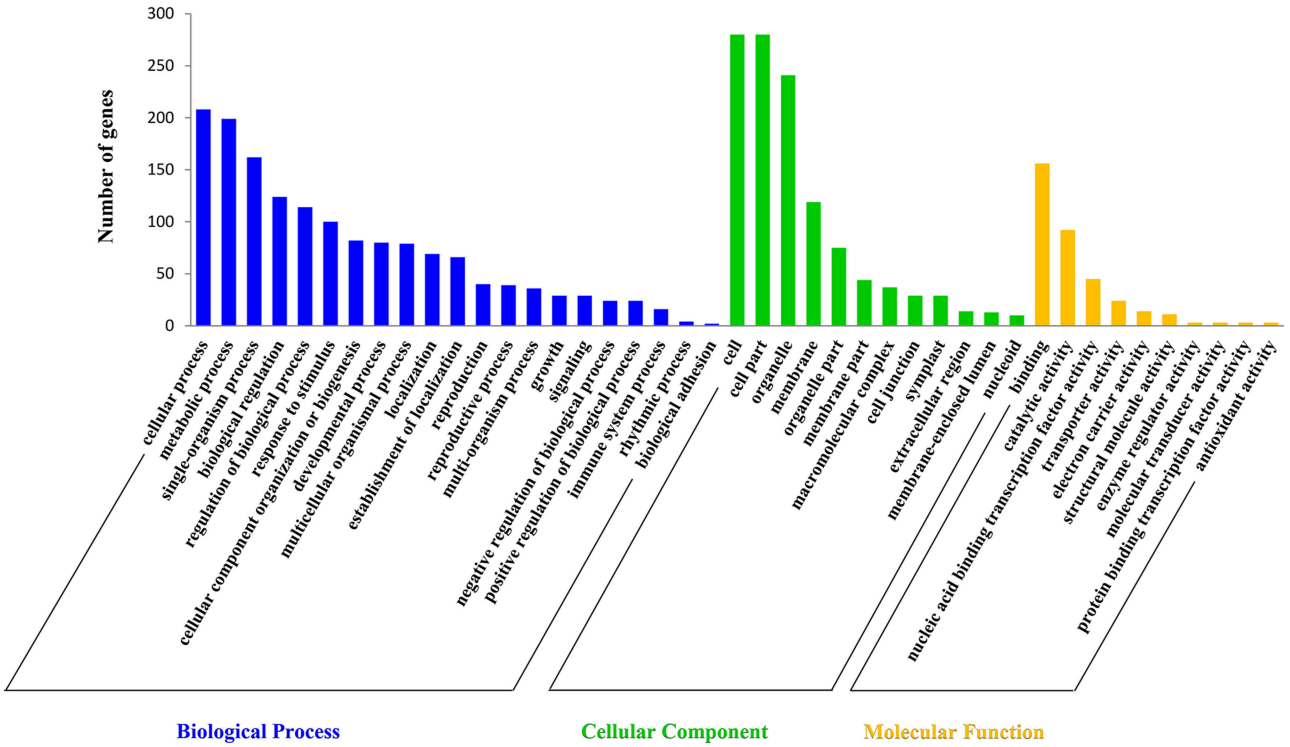


FIGURE 3 | Gene ontology classification of the predicted targets for differentially expressed miRNAs.

qRT–PCR Validation of miRNAs and Their Targets during Anther Development

To verify the quality of small RNA sequencing and analyze the expression patterns of CMS occurrence-related miRNAs in radish, a total of 15 miRNAs were randomly selected for qRT-PCR analysis. It was shown that the expression patterns of these miRNAs from qRT-PCR displayed a similar tendency with those from small RNA sequencing (**Figure 4**). To further study the dynamic expression patterns of CMS occurrence-related miRNAs and their corresponding targets during anther development, a total of 12 predicted target genes, *SPL3* (Rsa#S43017568 targeted by miR156a), *PPR* (Rsa#S42049270 targeted by miR158b-3p), *ARF16* (Rsa#S42581764 targeted by miR160a), *HRE1* (Rsa#S43010415 targeted by miR164b-3P), *TIR1* (FD955493 targeted by miR393a), *AGO5* (Unigene20881 targeted by miR396b), *Transducin/WD-40* (Rsa#S41989522 targeted by miR396b-3p), *F-box* (CL2205.Contig1 targeted by miR3444a-5p), *HB20* (Rsa#S43028702 targeted by rsa-miRn13), *NAC096* (Unigene22510 targeted by rsa-miRn15), *RDM4* (CL8993.Contig1 targeted by rsa-miRn17), and *UBQ1* (Rsa#S42012413 targeted by rsa-miRn27), were examined by qRT-PCR at three different stages of meiosis, tetrad, and early microspore. As shown in **Figure 5**, miR158b-3p, miR160a, miR164b-3p, and miR396b-3p were up-regulated and the expression levels maximized at meiosis stage, and then decreased at tetrad and early microspore stage. In addition, miR156a, miR393a, and miR3444a-5p were down-regulated at meiosis stage, and the expression levels then peaked at tetrad stage, but rapidly decreased at early microspore stage. miR396b showed an up-regulated expression pattern and peaked at tetrad stage, and then slightly decreased at early microspore stage. For the novel miRNAs, the expression levels of rsa-miRn13 and rsa-miRn27 were up-regulated at meiosis and tetrad stage, but dramatically decreased to the minimum at early microspore stage. Moreover, rsa-miRn15 was down-regulated at meiosis and tetrad stage, but rapidly increased to the maximum at early microspore stage. Transcripts of rsa-miRn17 reached its maximum at meiosis stage, but sharply declined at tetrad and early microspore stage. Furthermore, some negative correlations could be found between the expression levels of miRNAs and their corresponding target genes during various anther development stages, suggesting that miRNA-mediated mRNA silencing may be involved in CMS occurrence during anther development in 'WA' and 'WB' line (**Figure 5**).

DISCUSSION

High-throughput sequencing technology helps identify a large number of miRNAs and targets associated with CMS occurrence during anther development in several plant species (Wei et al., 2011, 2013; Fang et al., 2014; Yan et al., 2015), and provide an effective way to evaluate the expression profiles of miRNAs and targets in different tissues at different developmental stages. The production of functional pollen grain is a prerequisite

TABLE 4 | Identified candidate targets for some known and potential novel miRNAs during anther development.

miRNA	Target sequence	Target gene	Target gene annotation
miR156a	Rsa#S41982434	*CYP705A15*	Cytochrome P450, family 705, subfamily A, polypeptide 15
	Rsa#S43017568	*SPL3*	Squamosa promoter-binding-like protein 3
	CL289.Contig1	*SPL5*	Squamosa promoter-binding-like protein 5
	CL2234.Contig1	*SPL13*	Squamosa promoter-binding-like protein 13
	Unigene3780	*OTP82*	Chloroplast RNA editing factor
miR158b-3p	Rsa#S42049270	*PPR*	Pentatricopeptide repeat-containing protein
miR159a	Rsa#S42037487	*MYB101*	Myb domain protein 101
	Rsa#S41979156	*SPL*	Putative transcription factor SPL
	CL8717.Contig1	*SPL*	Putative transcription factor SPL
miR160a	Rsa#S42581764	*ARF16*	Auxin response factor 16
	Unigene466		Hydroxymethylglutaryl-CoA lyase
miR161	CL1282.Contig1	*MSL10*	Mechanosensitive channel of small conductance-like 10
	Unigene28541		Transcription initiation factor TFIIE alpha subunit
miR169b	CL2169.Contig1		26S proteasome non-ATPase regulatory subunit 14
miR169m	CL8331.Contig1		Sulfite exporter TauE/SafE family protein
miR171a-3p	FD953436		Peroxisomal nicotinamide adenine dinucleotide carrier
	CL271.Contig2	*LTP4*	Non-specific lipid-transfer protein 4
miR393a	FD955493	*TIR1*	Protein TRANSPORT INHIBITOR RESPONSE 1
	Unigene359	*EMB2726*	Elongation factor Ts family protein
miR395a	Unigene14836		Putative F-box/kelch-repeat protein
miR396a	CL879.Contig1		Transducin/WD-40 repeat-containing protein
	CL6202.Contig1	*SIP2;1*	Putative aquaporin SIP2-1
miR396b	Unigene20881	*AGO5*	Argonaute 5
	Unigene22800	*PPR*	Pentatricopeptide repeat-containing protein
miR396b-3p	Rsa#S41989522		Transducin/WD-40 repeat-containing protein
miR397a	CL379.Contig2		D-glycerate 3-kinase
	Unigene14031		Syntaxin/t-SNARE family protein
miR403	Rsa#S41987411	*AGO2*	Argonaute 2
	CL3585.Contig3	*AGO2*	Argonaute 2
miR482c-5p	CL561.Contig2		Serine/threonine protein kinase
	CL561.Contig4		Protein kinase family protein
miR854	Rsa#S42041817		Carboxylate clamp-tetratricopeptide repeat protein HOP2
	Rsa#S41978503	*F-box*	F-box protein
	CL5880.Contig1	*bZIP*	bZIP transcription factor
miR1878-3p	Rsa#S42043459		Putative metal tolerance protein C3
miR1885b	CL9579.Contig1	*CSLE1*	Cellulose synthase-like protein E1
	Unigene1615		Probable 26S proteasome non-ATPase regulatory subunit 3b
miR2111a-3p	CL444.Contig1	*AHL19*	AT-hook motif nuclear-localized protein

(Continued)

TABLE 4 | Continued

miRNA	Target sequence	Target gene	Target gene annotation
miR2111b-3p	CL4779.Contig1	*GLN1.3*	Glutamine synthetase cytosolic isozyme 1-3
miR2199	Rsa#S42571099	*AGL16*	Agamous-like MADS-box protein AGL16
miR3444a-5p	Rsa#S42004586	*PRX Q*	Peroxiredoxin Q
	Rsa#S42563276	*AGP16*	Arabinogalactan protein 16
	Rsa#S43011644	*FLA3*	Fasciclin-like arabinogalactan protein 3
	CL2205.Contig1	*F-box*	F-box protein
rsa-miRn2	Rsa#S42571626		Putative pectate lyase 18
rsa-miRn10	CL6997.Contig1	*GR-RBP2*	Glycine-rich RNA-binding protein 2
	CL7005.Contig1	*ATR1*	NADPH–cytochrome P450 reductase 1
rsa-miRn13	FD957134	*HB20*	Homeobox-leucine zipper protein ATHB-20
rsa-miRn15	Unigene22510	*NAC096*	NAC domain containing protein 96
rsa-miRn16	Unigene25057		Bromo-adjacent homology (BAH) domain-containing protein
rsa-miRn20	CL9688.Contig1	*PKP-BETA1*	Plastidial pyruvate kinase 2
	Unigene18839		Anticodon-binding domain-containing protein

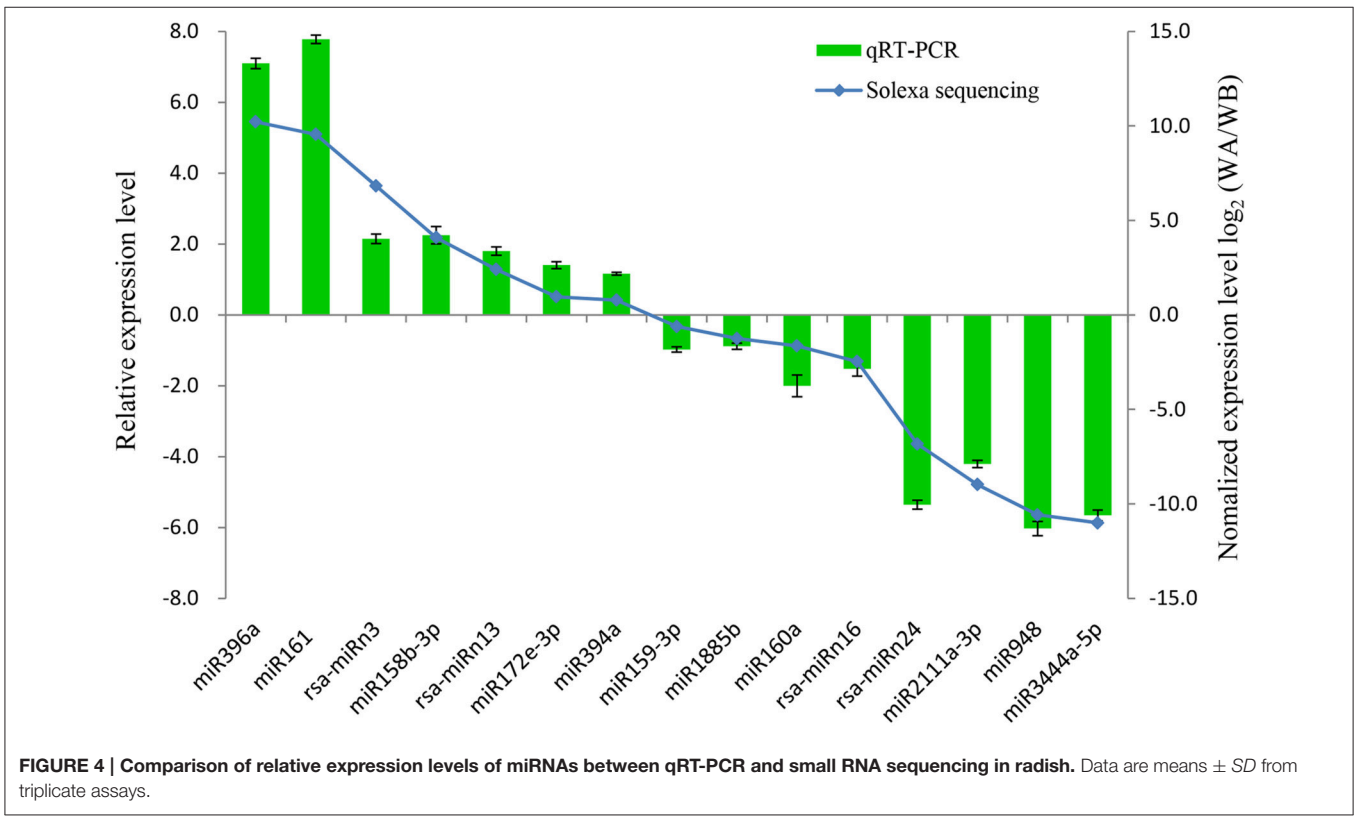


FIGURE 4 | Comparison of relative expression levels of miRNAs between qRT-PCR and small RNA sequencing in radish. Data are means ± *SD* from triplicate assays.

for the propagation in flowering plants, and the tapetum cell plays a critical role in microspore and pollen formation (Goetz et al., 2001). Unlike the radish CMS line 'WA' having no pollen in aborted anthers, its maintainer line 'WB' has normal anthers with fertile pollen (**Figure S1**). Cytological studies show that there is no visible difference between these

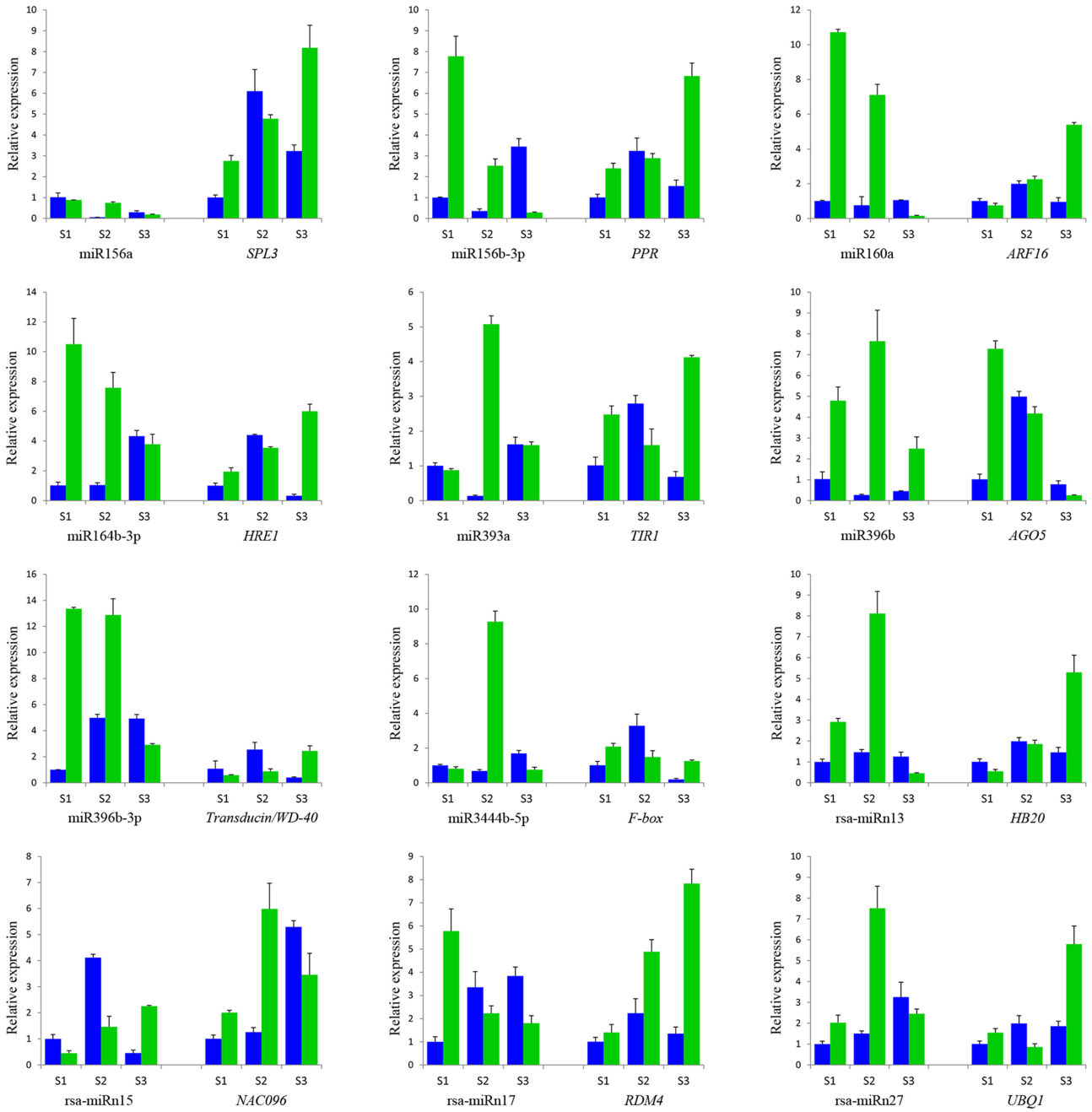


FIGURE 5 | qRT-PCR validation of differentially expressed miRNAs and its corresponding target genes in floral buds of 'WA' (in green) and 'WB' (in blue). (S1) meiosis stage, (S2) tetrad stage, (S3) early microspore stage. Data are means ± *SD* from triplicate assays.

two lines during the meiosis and tetrad stage (**Figure S1**). Thereafter, as compared with 'WB', the expanded, and vacuolated tapetum cells of 'WA' resulted in microspore degeneration and finally aborted anther with no pollen grains (**Figure S1**). However, few studies on the relationships between miRNAs and CMS occurrence during anther development in radish were conducted. The lack of CMS occurrence-related genes seriously hampered our understanding of molecular mechanism in CMS occurrence, which became an obstacle to utilize the heterosis of radish. To uncover the miRNA-mediated regulatory network of CMS occurrence during anther development, a comparative small RNAome sequencing was conducted in 'WA' and 'WB' line in this study. To our current knowledge, this study is the first investigation on identification and characterization of miRNAs, and their targets during anther development in radish.

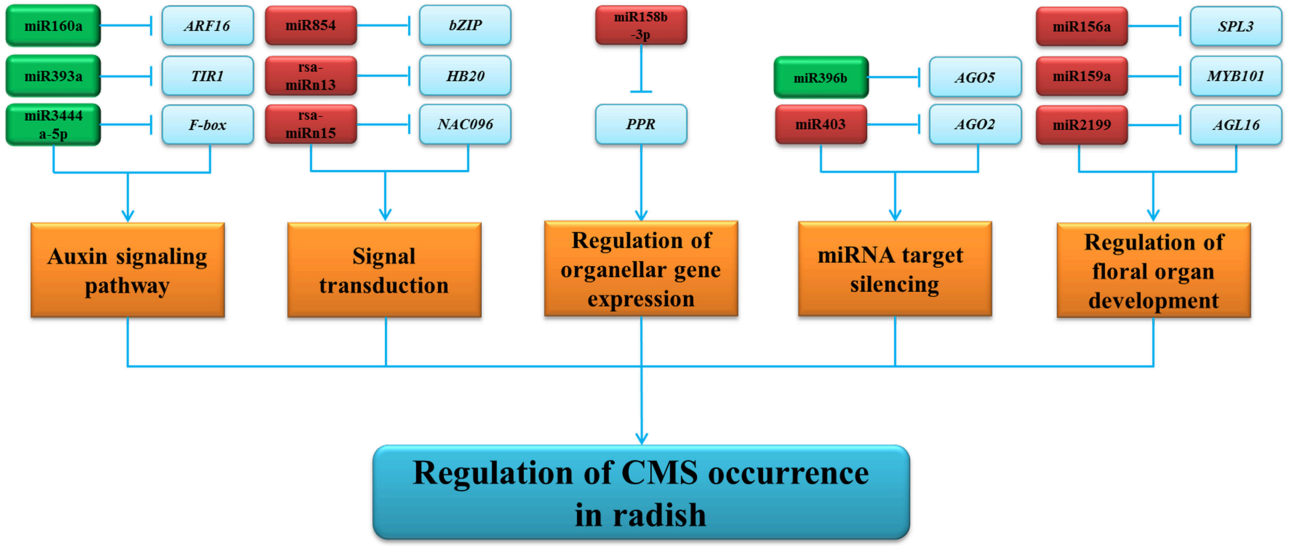


FIGURE 6 | The hypothetical schematic model of miRNA-mediated regulatory network of CMS occurrence during anther development in radish. The up- and down-regulated miRNAs are in red and green, respectively. Agamous-like MADS-box protein AGL16 (*AGL16*), argonaute 2 (*AGO2*), argonaute 5 (*AGO5*), auxin response factor 16 (*ARF16*), basic leucine-zipper (*bZIP*), F-box protein (*F-box*), myb domain protein 101 (*MYB101*), NAC domain containing protein 96 (*NAC096*), pentatricopeptide repeat-containing protein (*PPR*), protein TRANSPORT INHIBITOR RESPONSE 1 (*TIR1*), squamosa promoter-binding like protein 3 (*SPL3*).

With the application of high-throughput sequencing technology, it has provided an efficient tool to identify a quite comprehensive set of miRNAs at different stages and to reveal the miRNA-mediated regulatory network of CMS occurrence during anther development in plant. In this study, the length distribution of sRNAs suggested that the 24 nt sRNAs were the most abundant, followed by 21 nt sRNAs, which has been reported in *Arabidopsis* (Voinnet, 2009), *Prunus mume* (Gao et al., 2012), *O. sativa* (Ma et al., 2013), and *Medicago truncatula* (Eyles et al., 2013). The whole frequent percent of 21 and 24 nt small RNAs (28.33 and 30.07%, respectively) in radish was strikingly different from that of *B. juncea*, which 21 nt RNAs had high abundance (> 95%), and 24 nt RNAs possessed low frequency (1.1%) (Yang et al., 2013). Interestingly, the same tendency also existed when compared with *B. rapa* in which 24 nt sRNAs were the most dominant, followed by 21, 22, and 23 nt small RNAs (Jiang et al., 2014), it could be speculated that the genetic relationship between radish and *B. rapa* is closer than that between radish and *B. juncea* in the process of evolution.

Identification a set of miRNAs is a crucial step to promote our understanding of miRNA-mediated regulatory network of anther development and CMS occurrence. Recently, numerous studies have presented that the majority of known miRNAs in plantae are evolutionarily conserved (Chen et al., 2012; Barvkar et al., 2013). The diversity of known miRNA families in radish might be decided by the abundance and number of members. In the present study, a large number of conserved miRNAs expressed relatively higher levels compared with non-conserved ones, which was in agreement with previous researches in other species (Gao et al., 2012; Wang F. D. et al., 2012; Wang Z. J. et al., 2012; Fang et al., 2014). In addition, several studies have reported a number of known and potential novel miRNAs involved in anther development and CMS occurrence in *B. juncea* (Yang et al., 2013), *B. rapa* (Jiang et al., 2014), *Citrus reticulata* (Fang et al., 2014), *G. hirsutum* (Wei et al., 2013), and *O. sativa* (Yan et al., 2015), which greatly enhanced our knowledge of the regulatory roles of miRNAs in CMS occurrence. In this study, 28 known miRNAs were differentially expressed and the majority of these miRNAs were down-regulated during anther development. The differential expression patterns of rsa-miR160a and rsa-miR169b were consistent with those observed in *O. sativa* (Yan et al., 2015). Moreover, the expression pattern of rsa-miR396b and rsa-miR171a-3p was similar to that identified in *G. hirsutum* and *B. rapa*, respectively (Wei et al., 2013; Jiang et al., 2014). Interestingly, the targets of the miR160 contain three critical regulators, *ARF10*, *ARF16,* and *ARF17*, which are important in mediating gene expression response to the plant hormone auxin and regulating floral organ formation (Mallory et al., 2005; Wang et al., 2005; Chapman and Estelle, 2009; Liu et al., 2010). The expression level of rsa-miR160a was down-regulated in 'WA' and validated by qRT-PCR (**Figures 5**, **S5**). Thus, it could be speculated that the decreased abundance of rsa-miR160a may partially increase the expression of *ARF16*, finally resulting in abnormal pollen development in the sterile line 'WA'.

According to the negative correlation between differentially expressed miRNAs and their corresponding targets (**Figure S5**), a hypothetical schematic model of miRNA-mediated regulatory network of CMS occurrence during anther development in radish was put forward (**Figure 6**). As shown in the regulatory network, targets of these differentially expressed miRNAs containing

important transcription factors (TFs) and functional proteins are involved in many biological processes, including auxin signaling pathways, signal transduction, miRNA target silencing, floral organ development, and organellar gene expression. For instance, SBP-box genes targeted by miR156, a group of TFs with significant regulatory functions controlling the transition from the vegetative phase to the floral phase in *Arabidopsis*, *O. sativa*, and *Zea mays* (Chuck et al., 2007; Gandikota et al., 2007; Jiao et al., 2010). It was reported that three genes, *LEAFY*, *FRUITFULL,* and *APETALA1*, are directly activated by *SPL3* to regulate the timing of flower formation (Yamaguchi et al., 2009). Additionally, multiple SPL genes can lead to fully fertile flowers and regulate cell division and differentiation in *Arabidopsis* (Xing et al., 2010). In the present study, up-regulation of the rsa-miR156a decreased the expression of *SPL3* in 'WA' compared to 'WB' (**Figure 6**, **Table S7**), leading to disordered floral organ development, cell division, and differentiation in radish. MiR159 is required for normal anther development, in which it regulates the expression of genes encoding MYB TFs (Achard et al., 2004; Tsuji et al., 2006). MYB TFs are involved in the control of plant development, determination of cell fate and identity, primary, and secondary metabolism (Stracke et al., 2007; Gonzalez et al., 2008; Kang et al., 2009). *AtMYB103*, specifically expressed in tapetums and middle layers of anthers, is important for pollen development, especially the pollen exine formation (Zhang et al., 2007; Chen et al., 2014). Down-regulation of the *AtMYB103* resulted in earlier degeneration of tapetum and pollen grains aberration during anther development in *A. thaliana* (Zhang et al., 2007). In rice, anther and pollen defect in floral organ development are also found in the loss-of-function mutations of MYB (Kaneko et al., 2004). In the present study, the rsa-miR159a was found to be up-regulated in 'WA' line compared to 'WB' line (**Figures 6**, **S5**), indicating that the increased abundance of rsa-miR159a partially decreased the expression of *MYB101*, hampering normal tapetum development, callose dissolution, and exine formation in radish anthers. Moreover, *AGL16*, belonging to MADS-box transcription factors, was identified to be targeted by rsa-miR2199. The MADS-box TFs are essential regulators of the development of the floral meristems and floral organs in plants (Chen et al., 2014). These evidences indicated that rsa-miR2199 might be an essential component of gene regulatory network that involved in radish CMS occurrence (**Figure 6**).

Apart from key TFs, a variety of genes which encode important functional proteins, such as PPR proteins, F-box proteins, AGO proteins, and protein TRANSPORT INHIBITOR RESPONSE 1 (TIR1), were also considered to play important roles in CMS occurrence during anther development. PPR protein genes were identified as targets of miR158 (Lurin et al., 2004; Sunkar and Zhu, 2004). Previous studies indicated that PPR proteins are mostly located in the mitochondria and chloroplast and play crucial roles in pollen development, specific RNA sequence binding, post-transcriptional splicing and mRNA stability regulating (Okuda et al., 2006; Wang et al., 2006; Saze and Kakutani, 2007; Fujii and Small, 2011). In addition, some PPR proteins have also been identified as fertility-restoring genes (*Rf*) for CMS occurrence (Desloire et al., 2003; Wang et al., 2008; Yasumoto et al., 2009). In this study, rsa-miR158b-3p targeting the gene encoding PPR protein was up-regulated and the *PPR* gene was suppressed in 'WA' line compared with 'WB' line, and it could be suggested that the regular expression of CMS-associated mitochondrial genes and suppression of *PPR* gene result in sterility in radish 'WA' line (**Figures 5**, **S5**). Moreover, F-box proteins are involved in the regulation of various developmental processes in plants, including floral meristem, floral organ identity determination, and photomorphogenesis (Jain et al., 2007). The expression of rsa-miR3444a-5p was down-regulated at meiosis stage, and then peaked at tetrad stage, but rapidly decreased at early microspore stage, and a negative correlation was found between the expression levels of rsa-miR3444a-5p and its target gene which encoding F-box protein at three different stages according to the qRT-PCR analysis (**Figure 5**). In addition, *F-box* gene targeted by osa-miR528 was found to be involved in the regulation of the abortion process in male sterile line of rice. Moreover, the other 23 genes including *APG2*, *AGP16*, *FIO1*, *FLA3*, *FLA5*, *NAC083*, *NSP5*, *TRP1*, and *VIP1* were also the targets of rsa-miR3444a-5p, indicating that the miRNA has multiple effects on the targets (**Figure S5**). All of these genes targeted by rsa-miR3444a-5p might function together to regulate the CMS occurrence during anther development in radish. Additionally, AGO proteins were reported to be involved in diverse biological processes including hormone response, developmental regulation, and stress adaptation (Yang et al., 2013). Up-regulation of *TIR1* enhances auxin sensitivity, and causes altered leave phenotype and delayed flowering (Chen et al., 2011). In this study, *AGO2* and *AGO5* was targeted by miR403 and miR396b, respectively, and *TIR1* was targeted by miR393a, indicating miR403, miR396b, and miR393a might modulate the hormone response to play roles in the microspore development and CMS occurrence.

In summary, CMS occurrence-associated miRNAs and their targets between the male sterile line 'WA' and its maintainer line 'WB' were firstly identified and characterized in radish. These results provide a valuable foundation for unraveling the complex miRNA-mediated regulatory network of CMS occurrence and facilitate further dissection of roles of miRNAs during CMS occurrence and microspore formation in radish and other crops.

AUTHOR CONTRIBUTIONS

WZ, YX, and LL designed the research. WZ and YX conducted experiments. LX and YW participated in the design of the study and performed the statistical analysis. WZ and YX analyzed data and wrote the manuscript. XZ, RW, YZ, and EM helped with the revision of manuscript. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

This work was in part supported by grants from the National Key Technology R&D Program of China (2016YFD0100204-25), Key Technology R&D Program of Jiangsu Province (BE2016379), Jiangsu Agricultural Science and Technology Innovation Fund [JASTIF, CX(16)1012] and the PAPD.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01054

Figure S1 | Micrographs of anthers at different developmental stages in the CMS line 'WA' (A–E) and its maintainer line 'WB' (F–J). Panels **(A,F)** Meiosis stage. Panels **(B,G)** Tetrad stage. Panels **(C,H)** Early microspore stage. Panels **(D,I)** Pollen stage. Panels **(E,J)** Flower morphology.

Figure S2 | Distribution of known miRNA family members identified in radish.

Figure S3 | Abundance of each known miRNA family in radish.

Figure S4 | Precursor sequences and the predicted second structures of novel miRNAs in radish. The mature miRNAs are in red and miRNA*s are in blue (".” represent base mismatches, "(" represent base matches).

Figure S5 | The miRNA mediated regulatory network constructed by Cytoscape_v3.2.1. The red, yellow and green ellipses represent the know miRNAs, potential novel miRNAs and target genes, respectively.

Table S1 | Primers of miRNAs and targets in radish for qRT-PCR.

Table S2 | Statistical analysis of sequencing reads from the WA and WB sRNA library in radish.

Table S3 | Summary of common and specific sequences between WA and WB sRNA library.

Table S4 | Detailed information of known miRNAs identified from radish WA and WB library.

Table S5 | Detailed information of novel miRNAs identified from radish WA and WB library.

Table S6 | Differentially-expressed miRNAs between WA and WB in radish.

Table S7 | Putative targets of known and novel miRNAs identified in radish.

Table S8 | Predicted targets for non-conserved miRNAs in radish.

Table S9 | The detailed information of miRNA-targets for regulatory network construction.

REFERENCES

Achard, P., Herr, A., Baulcombe, D. C., and Harberd, N. P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development* 131, 3357–3365. doi: 10.1242/dev.01206

Allen, E., Xie, Z., Gustafson, A. M., and Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121, 207–221. doi: 10.1016/j.cell.2005.04.004

Barvkar, V. T., Pardeshi, V. C., Kale, S. M., Qiu, S., Rollins, M., Datla, R., et al. (2013). Genome-wide identification and characterization of microRNA genes and their targets in flax (*Linum usitatissimum*). *Planta* 237, 1149–1161. doi: 10.1007/s00425-012-1833-5

Bentolila, S., Alfonso, A. A., and Hanson, M. R. (2002). A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10887–10892. doi: 10.1073/pnas.102301599

Bodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., et al. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. *Science* 320, 1185–1190. doi: 10.1126/science.1159151

Brown, G. G., Formanova, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., et al. (2003). The radish Rfo restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. *Plant J.* 35, 262–272. doi: 10.1046/j.1365-313X.2003.01799.x

Chambers, C., and Shuai, B. (2009). Profiling microRNA expression in Arabidopsis pollen using microRNA array and real-time PCR. *BMC Plant Biol.* 9:87. doi: 10.1186/1471-2229-9-87

Chapman, E. J., and Estelle, M. (2009). Mechanism of auxin-regulated gene expression in plants. *Annu. Rev. Genet.* 43, 265–285. doi: 10.1146/annurev-genet-102108-134148

Chen, L., and Liu, Y. G. (2014). Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* 65, 579–606. doi: 10.1146/annurev-arplant-050213-040119

Chen, L., Ren, Y., Zhang, Y., Xu, J., Sun, F., Zhang, Z., et al. (2012). Genome-wide identification and expression analysis of heat-responsive and novel microRNAs in *Populus tomentosa*. *Gene* 504, 160–165. doi: 10.1016/j.gene.2012.05.034

Chen, P., Ran, S. M., Li, R., Huang, Z. P., Qian, J. H., Yu, M. L., et al. (2014). Transcriptome de novo assembly and differentially expressed genes related to cytoplasmic male sterility in kenaf (*Hibiscus cannabinus* L.). *Mol. Breed.* 34, 1879–1891. doi: 10.1007/s11105-014-0702-8

Chen, Z. H., Bao, M. L., Sun, Y. Z., Yang, Y. J., Xu, X. H., Wang, J. H., et al. (2011). Regulation of auxin response by miR393-targeted transport inhibitor response protein 1 is involved in normal development in Arabidopsis. *Plant Mol. Biol.* 77, 619–629. doi: 10.1007/s11103-011-9838-1

Chuck, G., Cigan, A. M., Saeteurn, K., and Hake, S. (2007). The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. *Nat. Genet.* 39, 544–549. doi: 10.1038/ng2001

Dai, X., and Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 39, W155–W159. doi: 10.1093/nar/gkr319

Desloire, S., Gherbi, H., Laloui, W., Marhadour, S., Clouet, V., Cattolico, L., et al. (2003). Identification of the fertility restoration locus, Rfo, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep.* 4, 588–594. doi: 10.1038/sj.embor.embor848

Eckardt, N. A. (2006). Cytoplasmic male sterility and fertility restoration. *Plant Cell* 18, 515–517. doi: 10.1105/tpc.106.041830

Eyles, R. P., Williams, P. H., Ohms, S. J., Weiller, G. F., Ogilvie, H. A., Djordjevic, M. A., et al. (2013). microRNA profiling of root tissues and root forming explant cultures in *Medicago truncatula*. *Planta* 238, 91–105. doi: 10.1007/s00425-013-1871-7

Fang, Y. N., Qiu, W. M., Wang, Y., Wu, X. M., Xu, Q., and Guo, W. W. (2014). Identification of differentially expressed microRNAs from a male sterile *Ponkan mandarin* (*Citrus reticulata* Blanco) and its fertile wild type by small RNA and degradome sequencing. *Tree Genet. Genome* 10, 1567–1581. doi: 10.1007/s11295-014-0780-7

Fujii, S., and Small, I. (2011). The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytol.* 191, 37–47. doi: 10.1111/j.1469-8137.2011.03746.x

Gandikota, M., Birkenbihl, R. P., Höhmann, S., Cardon, G. H., Saedler, H., and Huijser, P. (2007). The miRNA156/157 recognition element in the 3′UTR of the Arabidopsis SBP box gene *SPL3* prevents early flowering by translational inhibition in seedlings. *Plant J.* 49, 683–693. doi: 10.1111/j.1365-313X.2006.02983.x

Gao, Z. H., Shi, T., Luo, X. Y., Zhang, Z., Zhuang, W. B., and Wang, L. J. (2012). High-throughput sequencing of small RNAs and analysis of differentially expressed microRNAs associated with pistil development in Japanese apricot. *BMC Genomics* 13:371. doi: 10.1186/1471-2164-13-371

Goetz, M., Godt, D. E., Guivarc'h, A., Kahmann, U., Chriqui, D., and Roitsch, T. (2001). Induction of male sterility in plants by metabolic engineering of the carbohydrate supply. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6522–6527. doi: 10.1073/pnas.091097998

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., et al. (2008). Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 44, 3–12. doi: 10.1016/j.ymeth.2007.09.009

Henderson, I. R., Zhang, X., Lu, C., Johnson, L., Meyers, B. C., Green, P. J., et al. (2006). Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nat. Genet.* 38, 721–725. doi: 10.1038/ng1804

Hu, J., Wang, K., Huang, W. C., Liu, G., Gao, Y., Wang, J. M., et al. (2012). The rice pentatricopeptide repeat protein RF5 restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine-rich protein GRP162. *Plant Cell* 24, 109–122. doi: 10.1105/tpc.111.093211

Jain, M., Nijhawan, A., Arora, R., Agarwal, P., Ray, S., Sharma, P., et al. (2007). F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 143, 1467–1483. doi: 10.1104/pp.106.091900

Jiang, J. X., Lv, M. L., Liang, Y., Ma, Z. M., and Cao, J. S. (2014). Identification of novel and conserved miRNAs involved in pollen development in *Brassica campestris* ssp. *chinensis* by high-throughput sequencing and degradome analysis. *BMC Genomics* 15:146. doi: 10.1186/1471-2164-15-146

Jiao, Y. Q., Wang, Y. H., Xue, D. W., Wang, J., Yan, M. X., and Liu, G. F. (2010). Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat. Genet.* 42, 541–544. doi: 10.1038/ng.591

Jones-Rhoades, M. W., Bartel, D. P., and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* 57, 19–53. doi: 10.1146/annurev.arplant.57.032905.105218

Kaneko, M., Inukai, Y., Ueguchi-Tanaka, M., Itoh, H., Izawa, T., Kobayashi, Y., et al. (2004). Loss-of-function mutations of the rice *GAMYB* gene impair α-amylase expression in aleurone and flower development. *Plant Cell* 16, 33–44. doi: 10.1105/tpc.017327

Kang, Y. H., Kirik, V., Hulskamp, M., Nam, K. H., Hagely, K., Lee, M. M., et al. (2009). The *MYB23* gene provides a positive feedback loop for cell fate specification in the Arabidopsis root epidermis. *Plant Cell* 21, 1080–1094 doi: 10.1105/tpc.108.063180

Kubo, T., Kitazaki, K., Matsunaga, M., Kagami, H., and Mikami, T. (2011). Male sterility-inducing mitochondrial genomes: how do they differ? *Crit. Rev. Plant Sci.* 30, 378–400. doi: 10.1080/07352689.2011.587727

Kurihara, Y., Takashi, Y., and Watanabe, Y. (2006). The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. *RNA* 12, 206–212. doi: 10.1261/rna.2146906

Li, R. Q., Yu, C., Li,. Y. R., Lam, T. W., Yiu, S. M., Kristiansen, K., et al. (2009). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25, 1966–1967. doi: 10.1093/bioinformatics/btp336

Liu, X. D., Huang, J., Wang, Y., Khanna, K., Xie, Z. X., Owen, H. A., et al. (2010). The role of floral organs in carpels, an Arabidopsis loss-of-function mutation in MicroRNA160a, in organogenesis and the mechanism regulating its expression. *Plant J.* 62, 416–428. doi: 10.1111/j.1365-313X.2010.04164.x

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lurin, C., Andrés, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyère, C., et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16, 2089–2103. doi: 10.1105/tpc.104.022236

Ma, X. X., Shao, C. G., Wang, H. Z., Jin, Y. F., and Meng, Y. J. (2013). Construction of small RNA-mediated gene regulatory networks in the roots of rice (*Oryza sativa*). *BMC Genomics* 14:510. doi: 10.1186/1471-2164-14-510

Mallory, A. C., Bartel, D. P., and Bartel, B. (2005). MicroRNA-directed regulation of Arabidopsis *AUXIN RESPONSE FACTOR17* is essential for proper development and modulates expression of early auxin response genes. *Plant Cell* 17, 1360–1375. doi: 10.1105/tpc.105.031716

Meyers, B. C., Axtell, M. J., Bartel, B., Bartel, D. P., Baulcombe, D., Bowman, J. L., et al. (2008). Criteria for annotation of plant MicroRNAs. *Plant Cell* 20, 3186–3190. doi: 10.1105/tpc.108.064311

Nie, S. S., Xu, L., Wang, Y., Huang, D. Q., Muleke, E. M., Sun, X. C., et al. (2015). Identification of bolting-related microRNAs and their targets reveals complex miRNA-mediated flowering-time regulatory networks in radish (*Raphanus sativus* L.). *Sci. Rep.* 5:14034. doi: 10.1038/srep14034

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T., and Shikanai, T. (2006). A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. *J. Biol. Chem.* 28, 37661–37667. doi: 10.1074/jbc.M608184200

Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev.* 20, 3407–3425. doi: 10.1101/gad.1476406

Ruiz-Ferrer, V., and Voinnet, O. (2009). Roles of plant small RNAs in biotic stress responses. *Annu. Rev. Plant Biol.* 60, 485–510. doi: 10.1146/annurev.arplant.043008.092111

Saze, H., and Kakutani, T. (2007). Heritable epigenetic mutation of a transposon-flanked Arabidopsis gene due to lack of the chromatin-remodeling factor DDM1. *EMBO J.* 26, 3641–3652. doi: 10.1038/sj.emboj.7601788

Schnable, P. S., and Wise, R. P. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci.* 3, 175–180. doi: 10.1016/S1360-1385(98)01235-7

Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431–432. doi: 10.1093/bioinformatics/btq675

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K., et al. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant J.* 50, 660–677. doi: 10.1111/j.1365-313X.2007.03078.x

Sun, X. J., Liu, Y., Wang, L. J., Zhu, X. W., Gong, Y. Q., Xu, L., et al. (2012). Molecular characterization of the *Rs-Rf1* gene and molecular marker-assistional development of elite radish (*Raphanus sativus* L.) CMS lines with a functional marker for fertility restoration. *Mol. Breeding* 30, 1727–1736. doi: 10.1007/s11032-012-9756-1

Sun, X. J., Wang, Y., Chen, Y. L., Xu, L., Jiang, L. N., Gong, Y. Q., et al. (2013). Differential gene expression profiling of Ogura CMS line and its maintainer in radish (*Raphanus sativus* L.). *Acta Physiol. Plant.* 35, 3413–3425. doi: 10.1007/s11738-013-1376-9

Sunkar, R., and Zhu, J. K. (2004). Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell* 16, 2001–2019. doi: 10.1105/tpc.104.022830

Touzet, P., and Meyer, E. H. (2014). Cytoplasmic male sterility and mitochondrial metabolism in plants. *Mitochondrion* 19, 166–171. doi: 10.1016/j.mito.2014.04.009

Tsuji, H., Aya, K., Ueguchi-Tanaka, M., Shimada, Y., Nakazono, M., Watanabe, R., et al. (2006). GAMYB controls different sets of genes and is differentially regulated by microRNA in aleurone cells and anthers. *Plant J.* 47, 427–444. doi: 10.1111/j.1365-313X.2006.02795.x

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell* 136, 669–687. doi: 10.1016/j.cell.2009.01.046

Wang, F. D., Li, L. B., Liu, L. F., Li, H. Y., Zhang, Y. H., Yao, Y. Y., et al. (2012). High-throughput sequencing discovery of conserved and novel microRNAs in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). *Mol. Genet. Genomics* 287, 555–563. doi: 10.1007/s00438-012-0699-3

Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., and Chen, X. Y. (2005). Control of root cap formation by microRNA-targeted auxin response factors in Arabidopsis. *Plant Cell* 17, 2204–2216. doi: 10.1105/tpc.105.033076

Wang, Y., Liu, W., Shen, H., Zhu, X. W., Zhai, L. L., Xu, L., et al. (2015). Identification of radish (*Raphanus sativus* L.) miRNAs and their target genes to explore miRNA-mediated regulatory networks in lead (Pb) stress responses by high-throughput sequencing and degradome analysis. *Plant Mol. Biol. Rep.* 33, 358–376. doi: 10.1007/s11105-014-0752-y

Wang, Z. H., Zou, Y. J., Li, X. Y., Zhang, Q. Y., Chen, L., Wu, H., et al. (2006). Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* 18, 676–687. doi: 10.1105/tpc.105

Wang, Z. J., Huang, J. Q., Huang, Y. J., Li, Z., and Zheng, B. S. (2012). Discovery and profiling of novel and conserved microRNAs during flower development in *Carya cathayensis* via deep sequencing. *Planta* 236, 613–621. doi: 10.1007/s00425-012-1634-x

Wang, Z. W., Zhang, Y. J., Xiang, C. P., Mei, S. Y., Zhou, Y., Chen, G. P., et al. (2008). A new fertility restorer locus linked closely to the Rfo locus for cytoplasmic male sterility in radish. *Theor. Appl. Genet.* 117, 313–320. doi: 10.1007/s00122-008-0776-5

Wei, L. Q., Yan, L. F., and Wang, T. (2011). Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. *Genome Biol.* 12:R53. doi: 10.1186/gb-2011-12-6-r53

Wei, M. M., Wei, H. L., Wu, M., Song, M. Z., Zhang, J. F., Yu, J. W., et al. (2013). Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton. *BMC Plant Biol.* 13:66. doi: 10.1186/1471-2229-13-66

Xie, C., Mao, X. Z., Huang, J. J., Ding, Y., Wu, J. M., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. doi: 10.1093/nar/gkr483

Xing, S. P., Salinas, M., Höhmann, S., Berndtgen, R., and Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. *Plant Cell* 22, 3935–3950. doi: 10.1105/tpc.110.079343

Xu, L., Wang, Y., Xu, Y. Y., Wang, L. J., Zhai, L. L., Zhu, X. W., et al. (2013b). Identification and characterization of novel and conserved microRNAs in radish (*Raphanus sativus* L.) using high-throughput sequencing. *Plant Sci.* 201, 108–114. doi: 10.1016/j.plantsci.2012.11.010

Xu, L., Wang, Y., Zhai, L. L., Xu, Y. Y., Wang, L. J., Zhu, X. W., et al. (2013a). Genome-wide identification and characterization of cadmium-responsive microRNAs and their target genes in radish (*Raphanus sativus* L.) roots. *J. Exp. Bot.* 64, 4271–4287. doi: 10.1093/jxb/ert240

Yamaguchi, A., Wu, M. F., Yang, L., Wu, G., Poethig, R. S., and Wagner, D. (2009). The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of *LEAFY*, *FRUITFULL* and *APETALA1*. *Dev. Cell* 17, 268–278. doi: 10.1016/j.devcel.2009.06.007

Yan, J. J., Zhang, H. Y., Zheng, Y. Z., and Ding, Y. (2015). Comparative expression profiling of miRNAs between the cytoplasmic male sterile line MeixiangA and its maintainer line MeixiangB during rice anther development. *Planta* 241, 109–123. doi: 10.1007/s00425-014-2167-2

Yang, J. H., Liu, X. Y., Xu, B. C., Zhao, N., Yang, X. D., and Zhang, M. F. (2013). Identification of miRNAs and their targets using high-throughput sequencing and degradome analysis in cytoplasmic male-sterile and its maintainer fertile lines of *Brassica juncea*. *BMC Genomics* 14:9. doi: 10.1186/1471-2164-14-9

Yasumoto, K., Terachi, T., and Yamagishi, H. (2009). A novel Rf gene controlling fertility restoration of Ogura male sterility by RNA processing of orf138 found in Japanese wild radish and its STS markers. *Genome* 52, 495–504. doi: 10.1139/G09-026

Yu, B., Yang, Z. Y., Li, J. J., Minakhina, S., Yang, M. C., Padgett, R. W., et al. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935. doi: 10.1126/science.1107130

Yu, R. G., Wang, Y., Xu, L., Zhu, X. W., Zhang, W., Wang, R. H., et al. (2015). Transcriptome profiling of root microRNAs reveals novel insights into taproot thickening in radish (*Raphanus sativus* L.). *BMC Plant Biol.* 15:30. doi: 10.1186/s12870-015-0427-3

Zhai, L. L., Xu, L., Wang, Y., Huang, D. Q., Yu, R. G., Limera, C., et al. (2014). Genome-wide identification of embryogenesis-associated microRNAs in radish (*Raphanus sativus* L.) by high-throughput sequencing. *Plant Mol. Biol. Rep.* 32, 900–915. doi: 10.1007/s11105-014-0700-x

Zhang, Z. B., Zhu, J., Gao, J. F., Wang, C., Li, H., Li, H., et al. (2007). Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. *Plant J.* 52, 528–538. doi: 10.1111/j.1365-313X.2007.03254.x

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415. doi: 10.1093/nar/gkg595

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhang, Xie, Xu, Wang, Zhu, Wang, Zhang, Muleke and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 May 2016
doi: 10.3389/fpls.2016.00682

Transcriptomic Analysis Identifies Differentially Expressed Genes (DEGs) Associated with Bolting and Flowering in Radish (*Raphanus sativus* L.)

*Shanshan Nie, Chao Li, Yan Wang, Liang Xu, Everlyne M. Muleke, Mingjia Tang, Xiaochuan Sun and Liwang Liu **

National Key Laboratory of Crop Genetics and Germplasm Enhancement, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (East China), Ministry of Agriculture of China, College of Horticulture, Nanjing Agricultural University, Nanjing, China

OPEN ACCESS

Edited by:
Sarvajeet Singh Gill, Maharshi Dayanand University, India

Reviewed by:
Aashish Ranjan, National Institute of Plant Genome Research, India
Krishna Kant Sharma, Maharshi Dayanand University, India
Narsingh Chauhan, Maharshi Dayanand University, India

***Correspondence:**
Liwang Liu
nauliulw@njau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *08 January 2016*
Accepted: *03 May 2016*
Published: *24 May 2016*

Citation:
Nie S, Li C, Wang Y, Xu L, Muleke EM, Tang M, Sun X and Liu L (2016) Transcriptomic Analysis Identifies Differentially Expressed Genes (DEGs) Associated with Bolting and Flowering in Radish (Raphanus sativus L.). Front. Plant Sci. 7:682. doi: 10.3389/fpls.2016.00682

The transition of vegetative growth to bolting and flowering is an important process in the life cycle of plants, which is determined by numerous genes forming an intricate network of bolting and flowering. However, no comprehensive identification and profiling of bolting and flowering-related genes have been carried out in radish. In this study, RNA-Seq technology was applied to analyze the differential gene expressions during the transition from vegetative stage to reproductive stage in radish. A total of 5922 differentially expressed genes (DEGs) including 779 up-regulated and 5143 down-regulated genes were isolated. Functional enrichment analysis suggested that some DEGs were involved in hormone signaling pathways and the transcriptional regulation of bolting and flowering. KEGG-based analysis identified 37 DEGs being involved in phytohormone signaling pathways. Moreover, 95 DEGs related to bolting and flowering were identified and integrated into various flowering pathways. Several critical genes including *FT*, *CO*, *SOC1*, *FLC,* and *LFY* were characterized and profiled by RT-qPCR analysis. Correlation analysis indicated that 24 miRNA-DEG pairs were involved in radish bolting and flowering. Finally, a miRNA-DEG-based schematic model of bolting and flowering regulatory network was proposed in radish. These outcomes provided significant insights into genetic control of radish bolting and flowering, and would facilitate unraveling molecular regulatory mechanism underlying bolting and flowering in root vegetable crops.

Keywords: ***Raphanus sativus* L., bolting and flowering, RNA-Seq, hormone signaling, differentially expressed genes (DEGs)**

INTRODUCTION

The developmental transition from vegetative growth to bolting and flowering is one of the most important traits in plant life cycle. Bolting and flowering time must be appropriately determined to ensure reproductive success under most favorable conditions (Amasino and Michaels, 2010; Srikanth and Schmid, 2011). Plants have evolved an intricate bolting and flowering genetic circuitry

in response to various endogenous and environmental signals including development, age, plant hormones, photoperiod, and temperature (Fornara et al., 2010; Capovilla et al., 2015; Kazan and Lyons, 2015). Molecular and genetic regulation of flowering has been extensively studied in the model plant *Arabidopsis thaliana*. Five major flowering pathways including vernalization, photoperiod, autonomous, aging and gibberellin (GA) pathways have been identified to govern bolting and flowering time (Amasino and Michaels, 2010; Fornara et al., 2010; Srikanth and Schmid, 2011), and a number of flowering-related genes involved in these pathways have been isolated and characterized in *Arabidopsis* (Fornara et al., 2010; Srikanth and Schmid, 2011).

The signals from flowering pathways converge on several floral pathway integrators such as *FLOWERING LOCUS T* (*FT*), *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*) and *LEAFY* (*LFY*), which are integrated into the genetic networks of flowering (Moon et al., 2005; Parcy, 2005). Among these integrators, the florigen gene *FT* is a central node of floral transition, whose transcriptional expression is positively regulated by *CONSTANS* (*CO*) encoding a putative zinc finger transcription factor (Suárez-López et al., 2001), while it is negatively regulated by *FLOWERING LOCUS C* (*FLC*), a flowering repressor encoding a MADS-box transcription factor (Michaels and Amasino, 1999). Different environmental factors affect plant flowering by modulating the expression of floral integrators and stimulating changes in plant hormone levels (Yaish et al., 2011; Riboni et al., 2014; Kazan and Lyons, 2015). Increasing evidences have revealed the connections between flowering time and plant hormones including salicylic acid (SA), jasmonic acid (JA), GA, abscisic acid (ABA) and auxin (Davis, 2009; Kazan and Lyons, 2015). The effects of phytohormone signaling on flowering, particularly GA pathway, have been extensively described in *Arabidopsis* (Mutasa-Göttgens and Hedden, 2009). Therefore, understanding the roles of flowering-related genes and crosstalk between diverse genetic pathways is fundamental for elucidating the regulatory mechanisms underlying bolting and flowering in plants.

RNA sequencing (RNA-Seq), a powerful strategy for global discovery of functional genes, has provided a better qualitative and quantitative description of gene expressions under certain conditions in many plant species (Lister et al., 2009; Wang et al., 2009a). Digital gene expression (DGE) tag profiling is a revolutionary approach for identifying differentially expressed genes (DEGs) in diverse plant tissues, organs and developmental stages (Bai et al., 2013; Zhang et al., 2014a; Zhu et al., 2015). Moreover, RNA-Seq combined with DGE profiling has been employed for flowering-related gene discovery and expression analysis in some species such as bamboo (Gao et al., 2014), *Lagerstroemia indica* (Zhang et al., 2014b), sweetpotato (Tao et al., 2013) and litchi (Zhang et al., 2014c). However, to our knowledge, there are no studies on global expression profile analysis of bolting and flowering-related genes in radish (*Raphanus sativus* L.).

Radish ($2n = 2x = 18$), belonging to Brassicaceae family, is an important annual or biennial root vegetable crop worldwide. Premature bolting is a seriously destructive problem and results in poor root growth and the reduced harvest during radish production, especially in spring. Appropriate timing of bolting and flowering is significant for reproductive success at suitable conditions, as well as preventing the premature bolting in radish. Progress on bolting and flowering time control (Fornara et al., 2010; Srikanth and Schmid, 2011), especially in *Arabidopsis*, has provided a solid foundation and reference for identifying numerous functional genes during radish bolting and flowering. Recently, the transcriptomes from radish roots and leaves have been assembled and analyzed (Wang et al., 2013; Zhang et al., 2013; Xu et al., 2015). Moreover, a list of microRNAs (miRNAs) and functional genes related to bolting and flowering were successfully isolated from late-bolting radish based on transcriptomic datasets (Nie et al., 2015). Therefore, to further identify the DEGs involved in bolting and flowering regulation is of importance for understanding the genetic regulatory network of bolting and flowering in radish.

In this study, to investigate the gene expression patterns during the transition of vegetative growth to bolting stage in radish, using the late bolting radish advanced inbred line as material, two DGE libraries were constructed and sequenced with RNA-Seq technology. The aims were to comprehensively identify DEGs involved in bolting and flowering regulatory network and to explore their roles in determining radish bolting and flowering time. Expression patterns of several critical DEGs related to bolting and flowering were validated by quantitative real-time PCR (RT-qPCR) analysis. Finally, to characterize the bolting and flowering-related genes and miRNAs in flowering pathways, a putative miRNA-DEG-based model of bolting and flowering regulatory network was put forward in radish. These results could provide significant insights into the molecular mechanism underlying bolting and flowering regulation in radish and other root vegetable crops.

MATERIALS AND METHODS

Plant Materials

The late bolting radish advanced inbred line 'NAU-LU127,' which was self-pollinated for more than 20 generations, was used in this study. The genetic background and structure of this line are stable and highly homozygous. After surface-sterilization, the seeds were sowed and grew in a growth chamber with 16 h light at 25°C and 8 h darkness at 16°C. The radish leaves used for DGE sequencing and RT-qPCR analysis were separately collected at two different developmental stages: vegetative stage (VS) and reproductive stage (RS), with three biological replicates. Each sample was collected at two developmental stages from three randomly selected individual plants, respectively. All the samples were immediately frozen in liquid nitrogen and stored at −80°C until use.

DGE Library Construction and Illumina Sequencing

Total RNA from radish leaves at vegetative stage and reproductive stage was individually extracted using Trizol® Reagent (Invitrogen) following the manufacturer's instructions. The equivalent quantity of total RNA from three replicates was pooled and used for library preparation and sequencing. Two

cDNA libraries named NAU-VS and NAU-RS were constructed and sequenced according to the previously reported method (Xu et al., 2015). The library construction and Illumina sequencing were performed using HiSeq™ 2500 platform at Beijing Genomics Institute (BGI, Shenzhen, China). The RNA-Seq data were deposited in NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra/) with accession numbers of SRX1671036 (NAU-VS) and SRX1671054 (NAU-RS).

Data Processing and Expression Analysis of DEGs

The raw reads were primarily produced for data processing. After filtering low quality reads, adaptor sequences and reads containing ploy-N, the clean reads were obtained. These clean reads were then matched to the radish reference sequences which contained the public radish genomic survey sequences (GSS) and expressed sequence tag (EST) sequences and leaf transcriptome sequences from 'NAU-LU127' (Nie et al., in press) with no more than two mismatches. These sequences from radish leaf transcriptome have been deposited in NCBI Transcriptome Shotgun Assembly (TSA, http://www.ncbi.nlm.nih.gov/genbank/tsa/) database under the accession number GEMG00000000.

To screen the DEGs between two DGE libraries, the expression level of each transcript is calculated using RPKM (Reads Per kb per Million reads) method (Mortazavi et al., 2008). Prior to differential gene expression analysis, the read counts of each transcript were adjusted by edgeR program package (Robinson et al., 2010) through one scaling normalized factor. Trimmed Mean of M values (TMM), an appropriate normalization method implemented in the edgeR package (Robinson and Oshlack, 2010; Robinson et al., 2010), was employed to obtain the normalized read counts. The differential expression analysis of two libraries was performed using the DEGSeq R package 1.20.0 (Wang et al., 2010). Subsequently, the false discovery rate (FDR) was used to determine *P*-value threshold in multiple testing (Benjamini et al., 2001). A strict algorithm was used to further perform DEG identification according to the previous reports (Audic and Claverie, 1997). The absolute value of $\log_2$Ratio (NAU-RS/NAU-VS) ≥ 1, $P < 0.05$ and FDR ≤ 0.001 were used as threshold for judging the significance of gene expression difference. The cluster analysis of gene expression patterns was performed with cluster software and Java Treeview software (Saldanha, 2004).

Functional Annotation and Enrichment Analysis of DEGs

To investigate the biological function and involvement in functional pathways, all the identified transcripts were mapped to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) database. For GO annotation, the unique transcripts were subjected to BLASTX searching against the NCBI Nr database using the $E < 10^{-5}$. Then the Blast2GO (Conesa et al., 2005) and WEGO software (Ye et al., 2006) were used to obtain GO annotations and functional classifications. GO enrichment analysis of DEGs was implemented by the GOseq R package (Young et al., 2010). KOBAS software (Xie et al., 2011) was used to test the statistical enrichment of DEGs in KEGG pathways. The significantly enriched functional terms and pathways were identified using the criterion of a Bonferroni-corrected $P \leq 0.05$.

RT-qPCR Analysis

Total RNAs from radish leaves were isolated and obtained as described above. RT-qPCR was performed on an iCycler Real-Time PCR Detection System (Bio-Rad, USA) with three replications according to previous reports (Nie et al., 2015; Xu et al., 2015). All PCR reactions were carried out in a total volume of 20 μL with *RsActin* gene as the internal control (Xu et al., 2012). The relative gene expression levels were calculated using $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001). The specific PCR primers were designed using Beacon Designer 7.0 (Premier Biosoft International, USA) and listed in **Table S7**.

RESULTS

DGE Library Sequencing and Data Analysis

To obtain global unique sequences from radish leaves, *de novo* assembly and analysis of transcriptome prepared from radish leaves of 'NAU-LU127' were carried out using Illumina RNA-Seq technology. Totally 111,167 contigs and 53,642 unigenes were generated from the radish leaf transcriptome (Nie et al., in press). The available dataset of radish leaf transcriptome integrating with the available radish GSS and EST sequences released in NCBI database enriched the radish reference sequences for DEG identification during radish bolting and flowering.

In this study, two DGE libraries from leaves of radish advanced inbred line 'NAU-LU127' at vegetative and reproductive stages were constructed and sequenced by Illumina HiSeq™ 2500 platform, respectively. As a result, 8,825,790 and 12,382,793 raw reads were obtained in NAU-VS and NAU-RS libraries, respectively (**Table 1**; **Figure 1A**). After removing adaptor sequences and low quality reads, 8,541,912 and 10,154,256 clean reads were generated in the two libraries (**Table 1**). These clean reads were then mapped to the radish reference sequences, resulting in the generation of 87.83% (7,502,455 reads) and 68.50% (6,955,386 reads) matched reads in NAU-VS and NAU-RS libraries, respectively. For the variation of clean reads mapping percentage between two libraries, it may arise from the sample differences and the specific pre-processing of obtained reads (Oshlack et al., 2010). The more mapped reads from NAU-VS library implied that some stage-specific genes may be expressed only at vegetative stage and differentially expressed between these two DGE libraries. Further analysis revealed that 4,124,632 reads (48.29 %) in NAU-VS library and 5,200,880 reads (51.22 %) in NAU-RS library were uniquely matched (**Table 1**).

Identification and Functional Enrichment Analysis of DEGs

The transcript abundance of each gene from two DGE libraries was calculated and analyzed by RPKM method. The threshold of $|\log_2\text{Ratio}| \geq 1$ and FDR ≤ 0.001 were further used to determine the significantly DEGs. A total of 5922 significantly DEGs

TABLE 1 | Summary of DGE sequencing and mapped reads.

Summary	NAU-VS		NAU-RS	
	Reads number	Percent (%)	Reads number	Percent (%)
Raw reads	8,825,790	100.00	12,382,793	100.00
Clean reads	8,541,912	100.00	10,154,256	100.00
Total base pairs	418,553,688	100.00	497,558,544	100.00
Total mapped reads	7,502,455	87.83	6,955,386	68.50
Perfect match	6,087,824	71.27	5,699,760	56.13
<=2bp mismatch	1,414,631	16.56	1,255,626	12.37
Unique match	4,124,632	48.29	5,200,880	51.22
Multi-position match	3,377,823	39.54	1,754,506	17.28
Unmapped reads	1,039,457	12.17	3,198,870	31.50

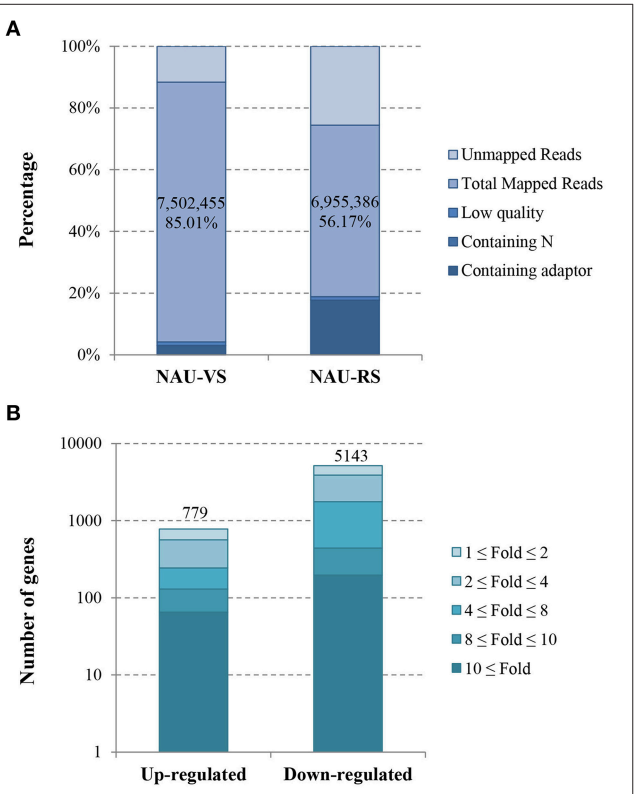


FIGURE 1 | Overview of the sequencing and mapping (A) and the identified DEGs (B) in NAU-VS and NAU-RS libraries.

including 779 up-regulated and 5143 down-regulated genes were obtained from NAU-VS and NAU-RS libraries (**Table S1**; **Figure 1B**).

To better classify the functions of these identified DEGs, GO classification and enrichment analysis were carried out in this study. These DEGs were categorized into three main GO categories including 23 biological processes, 14 cellular components and 13 molecular functions (**Figure 2**). Functional enrichment analysis revealed that 140 GO terms were significantly enriched with a Bonferroni-corrected $P \leq 0.05$ (**Table S2**). The terms of "metabolic process" (GO: 0008152) and "organic substance metabolic process" (GO: 0071704) were the dominant groups in biological processes; "cell" (GO: 0005623) and "cell part" (GO: 0044464) were the highly represented groups in the cellular components. For the molecular functions, a large proportion of genes were significantly enriched in "organic cyclic compound binding" (GO: 0097159) and "heterocyclic compound binding" (GO: 1901363) categories. Moreover, some enriched GO terms were related to plant flowering and meristem development, including "regulation of photoperiodism, flowering" (GO: 2000028), "regulation of timing of meristematic phase transition" (GO: 0048506), "meristem maintenance" (GO: 0010073), "meristem growth" (GO: 0035266), "meristem development" (GO: 0048507) and "flower development" (GO: 0009908) (**Table S2**).

To further understand the putative active biological pathways, all the identified DEGs were mapped to KEGG database by BLASTx with $E \leq 10^{-5}$ and $Q \leq 1$. As a result, 5922 DEGs were successfully assigned to 128 KEGG pathways (**Table S3**). The dominant pathway was "Metabolic pathways," followed by "Biosynthesis of secondary metabolites," "Ribosome," and "Plant hormone signal transduction." Moreover, 17 pathways were significantly enriched ($Q \leq 0.05$; **Table 2**), including "Circadian rhythm-plant" (ko04712), "Plant hormone signal transduction" (ko04075), "Photosynthesis" (ko00195), "Ribosome" (ko03010) and "Vitamin B6 metabolism" (ko00750).

DEGs Involved in Hormone Signal Transduction Pathway

In this study, pathway-based analysis showed that 37 DEGs representing 393 unique sequences were identified and involved in "Plant hormone signal transduction" (ko04075) pathway (**Table 3**; **Table S4**; **Figure S1**). These genes including *AUX1*, *TRANSPORT INHIBITOR RESPONSE 1* (*TIR1*), *AUXIN RESPONSE FACTOR* (*ARFs*), *GIBBERELLIN RECEPTOR 1* (*GID1*), and *CORONITINE INSENSITIVE 1* (*COI1*), participated in the regulation of several hormone homeostasis and flowering time (Davis, 2009; Kazan and Lyons, 2015). Enrichment analysis revealed that most of genes were involved in auxin, GA, ABA, JA, and SA signaling pathways (**Figure 3**). In GA signaling pathway, one down- and three up-regulated transcripts were related to GID1 protein, while eight down-regulated transcripts encoded DELLA protein (**Figure 3A**). For the process of JA signaling, two down-regulated transcripts encoded JAR1 protein and one down-regulated transcript encoded COI1 protein (**Figure 3D**). In auxin signaling pathway, 10 down-regulated transcripts belonged to *ARF* genes, while one up- and seven down-regulated transcripts encoded auxin-responsive proteins (**Figure 3E**). In addition, some DEGs related to other phytohormone biosynthesis were also identified, including zeatin biosynthesis (ko00908, three DEGs), carotenoid biosynthesis (ko00906, four DEGs), cysteine and methionine metabolism (ko00270, seven DEGs), brassinosteroid biosynthesis (ko00905, eight DEGs), and phenylalanine metabolism (ko00360, three DEGs) (**Table 3**; **Figure S1**).

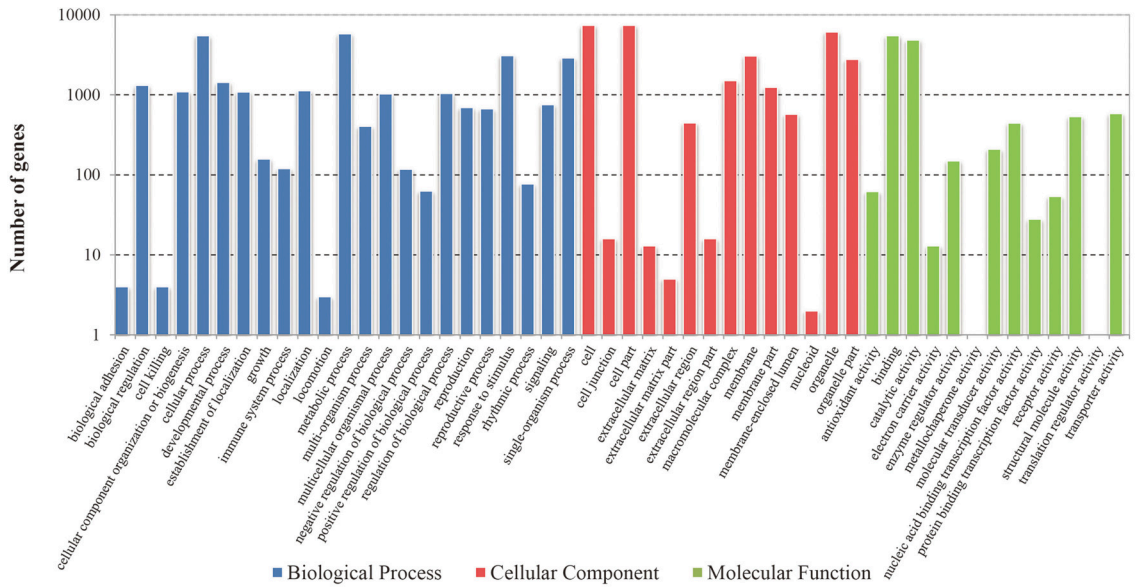


FIGURE 2 | The GO classification of identified DEGs in radish.

DEGs Involved in the Transition of Vegetative Growth to Bolting in Radish

In this study, to identify DEGs during radish bolting and flowering, BLAST searching was performed and the putative functions of DEGs were assessed. A total of 95 DEGs representing 128 unique sequences related to bolting and flowering were identified (**Table S5**). The analysis of flowering pathways revealed that these genes were involved in five different flowering pathways including photoperiod, vernalization, autonomous, GA and aging pathways.

In the present study, some unigenes representing photoperiodic flowering genes were identified, including *AGAMOUS-LIKE 24* (*AGL24*), *APETALA2* (*AP2*), *CO*, *CELL GROWTH DEFECT FACTOR 1* (*CDF1*), and *CONSTITUTIVELY PHOTOMORPHOGENIC 1* (*COP1*; **Table S5**). Some genes related to circadian rhythm and light signaling pathway included *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*), *CONSTANS-LIKE 1* (*COL1*), *CRYPTOCHROME* (*CRY2*), *TIMING OF CAB EXPRESSION 1* (*TOC1*), and *LATE ELONGATED HYPOCOTYL* (*LHY*; **Table S5**).

For the vernalization pathway, one down-regulated transcript (CL1584.Contig3) belonging to *FLC* homolog was found in this study (**Table S5**), which is a major flowering repressor and integrates the autonomous and vernalization pathways (Michaels and Amasino, 1999). Many genes involved in vernalization pathway including *FRIGIDA* (*FRI*), *FRIGIDA-like* (*FRL*), *FRIGIDA INTERACTING PROTEIN 2* (*FIP2*), *EMBRYONIC FLOWER 2* (*EMF2*), *VERNALIZATION 1* (*VRN1*), and *VERNALIZATION 2* (*VRN2*) were also identified and implicated in regulating the expression of *FLC*. Furthermore, *LUMINIDEPENDENS* (*LD*), *FPA*, *FVE,* and *FY* involved in autonomous pathway were also identified (**Table S5**).

Moreover, some putative genes for GA and aging pathways were also found in the present study (**Table S5**). The candidate genes involved in GA pathway comprised *GIGANTEA* (*GI*), *GNC*, *GA INSENSITIVE DWARF 1B* (*GID1B*), *DWARF AND DELAYED FLOWERING 1* (*DDF1*), and *REPRESSOR OF GA 1-3* (*RGA1-3*). The candidate genes related to aging pathway included *SQUAMOSA PROMOTER BINDING-LIKE PROTEIN 1* (*SPL1*), *SPL2*, *SPL3*, *SPL9*, *SPL13,* and *SPL15*. In addition, some floral integrators such as *FT* (FD571044), *SOC1* (CL4258.Contig1) and *LFY* (Unigene29702), were also identified in this study (**Table S5**).

Expression Profile Analysis by RT-qPCR

To validate the differential expression patterns of DEGs during radish bolting and flowering, totally 21 functional genes were randomly selected and subjected to RT-qPCR analysis. These selected genes included six genes related to hormone signaling and 15 genes related to bolting and flowering regulation. The relative expression levels of these genes between vegetative growth and reproductive stage were analyzed and compared (**Figure 4**). Further, comparative analysis revealed that these gene expression trends except *MYC2*-CL4584.Contig1 were in agreement with the transcript abundance changes by RNA-Seq (**Figure 4**), indicating the highly accuracy and quality of DGE sequencing.

The Regulatory Network Underlying Bolting and Flowering in Radish

Considerable studies have revealed that some miRNAs regulating corresponding target genes played important roles in the transition from vegetative growth to bolting and flowering (Spanudakis and Jackson, 2014). In our recent study, several bolting and flowering-related miRNA-target gene pairs were identified and characterized in late-bolting radish (Nie et al., 2015). To better understand the genetic regulatory network of radish bolting and flowering, correlation analysis between the

TABLE 2 | The significantly enriched pathways for DEGs in radish.

Pathway	Pathway ID	Number of DEGs	*P*-value	*Q*-value
Amino sugar and nucleotide sugar metabolism	ko00520	142 (1.89%)	1.65E-03	1.76E-02
Aminoacyl-tRNA biosynthesis	ko00970	71 (0.95%)	4.91E-06	2.09E-04
Carbon fixation in photosynthetic organisms	ko00710	104 (1.38%)	1.43E-03	1.67E-02
Circadian rhythm-plant	ko04712	101 (1.34%)	2.33E-03	1.50E-02
Histidine metabolism	ko00340	23 (0.31%)	8.10E-04	1.04E-02
Lysine biosynthesis	ko00300	24 (0.32%)	1.29E-04	3.32E-03
Other types of O-glycan biosynthesis	ko00514	13 (0.17%)	4.60E-03	4.20E-02
Pentose phosphate pathway	ko00030	56 (0.75%)	2.61E-04	3.83E-03
Photosynthesis	ko00195	115 (1.53%)	2.69E-04	3.83E-03
Photosynthesis-antenna proteins	ko00196	36 (0.48%)	6.50E-05	2.08E-03
Plant hormone signal transduction	ko04075	394 (5.25%)	7.14E-03	3.27E-02
Porphyrin and chlorophyll metabolism	ko00860	68 (0.91%)	2.56E-06	1.64E-04
Protein export	ko03060	64 (0.85%)	3.02E-03	2.98E-02
Ribosome	ko03010	504 (6.71%)	2.25E-04	3.83E-03
RNA degradation	ko03018	156 (2.08%)	3.70E-07	4.73E-05
RNA transport	ko03013	240 (3.2%)	2.62E-04	3.83E-03
Vitamin B6 metabolism	ko00750	16 (0.21%)	5.85E-03	4.99E-02

DEGs identified in the present study and bolting and flowering-related miRNAs previously reported (Nie et al., 2015) was performed. As expected, 24 miRNA-mRNA pairs including 16 miRNAs and 27 target DEGs were identified (**Table S6**). Among them, 19 miRNA-mRNA pairs showed negative correlations in expression patterns. Several DEGs including *AP2* (targeted by miR172), *VRN1* (targeted by miR5227), *PRP39* (targeted by miR6273), and *NF-YB3* (targeted by miR860), were found to be involved in bolting and flowering regulation (Wang et al., 2007; Kumimoto et al., 2008; Zhu and Helliwell, 2010).

To gain insights into the bolting and flowering regulatory network in radish, a putative model for summarizing the bolting and flowering-related DEGs and miRNAs was proposed (**Figure 5**). The critical genes involved in various flowering pathways and phytohormone signaling pathways were displayed in the schematic regulatory network of radish bolting and flowering. According to the known *Arabidopsis* flowering regulatory network (Fornara et al., 2010; Srikanth and Schmid, 2011), we speculated that the transcriptional regulations of several floral integrators including *FT*, *CO*, *SOC1*, *FLC,* and *LFY*, could integrate the signals from various pathways and modulate the radish bolting and flowering (**Figure 5**). Moreover, the models of miR172-*AP2* and miR5227-*VRN1* have been shown to be important participants in the regulatory network of bolting and flowering (Wang et al., 2007; Zhu and Helliwell, 2010; Nie et al., 2015).

TABLE 3 | The identified DEGs involved in hormone signal transduction pathway in radish.

Gene entry	Gene name	Signaling pathway	Annotation
K13946	*AUX1*	Auxin	Auxin influx carrier
K14485	*TIR1*	Auxin	Transport inhibitor response 1 TIR1
K14484	*AUX/IAA*	Auxin	Auxin-responsive protein IAA
K14486	*ARF*	Auxin	Auxin response factor
K14487	*GH3*	Auxin	Auxin-responsive GH3 family protein
K14488	*SAUR*	Auxin	SAUR-like auxin-responsive protein
K14489	*CRE1*	Cytokinin	Cytokinin receptor
K14491	*B-ARR*	Cytokinin	Two-component response regulator ARR-B family
K14492	*A-ARR*	Cytokinin	Two-component response regulator ARR-A family
K14493	*GID1*	Cytokinin	Gibberellin receptor 1
K14494	*DELLA*	Cytokinin	DELLA protein
K14496	*PYR/PYL*	Abscisic acid	Abscisic acid receptor PYL
K14497	*PP2C*	Abscisic acid	Probable protein phosphatase 2C
K14498	*SnRK2*	Abscisic acid	Serine-threonine-protein kinase SnRK2
K14432	*ABF*	Abscisic acid	Abscisic acid responsive element-binding factor
K14509	*ETR*	Ethylene	Ethylene response sensor
K14510	*CTR1*	Ethylene	Two-component response regulator ARR-A family
K14512	*MPK6*	Ethylene	Mitogen-activated protein kinase 6
K14513	*EIN2*	Ethylene	Ethylene-insensitive protein 2
K14515	*EBF1/2*	Ethylene	EIN3-binding F-box protein
K14514	*EIN3*	Ethylene	Ethylene-insensitive protein 3
K14517	*ERF1/2*	Ethylene	Ethylene-responsive transcription factor
K13416	*BAK1*	Brassinosteroids	Brassinosteroid insensitive 1-associated receptor kinase 1
K13415	*BRI1*	Brassinosteroids	Protein brassinosteroid insensitive 1
K14500	*BSK*	Brassinosteroids	BR-signaling kinase
K14501	*BSU1*	Brassinosteroids	BSU1 protein
K14502	*BIN2*	Brassinosteroids	Brassinosteroid insensitive protein2
K14503	*BZR1/2*	Brassinosteroids	BZR1 protein
K14504	*TCH4*	Brassinosteroids	Xyloglucan:xyloglucosyl transferase TCH4
K14505	*CYCD3*	Brassinosteroids	Cyclin D3, plant
K14506	*JAR1*	Jasmonate	JAR1 protein
K13463	*COI1*	Jasmonate	Coronitine insensitive 1
K13464	*JAZ*	Jasmonate	JAZ protein
K13422	*MYC2*	Jasmonate	bHLH transcription factor MYC2
K14508	*NPR1*	Salicylic acid	NPR1 protein
K14431	*TGA*	Salicylic acid	Transcription factor TGA
K13449	*PR-1*	Salicylic acid	Pathogenesis-related protein 1

DISCUSSION

Radish bolting and flowering are integral stages in its complete life cycle. The timing of bolting and flowering is coordinately

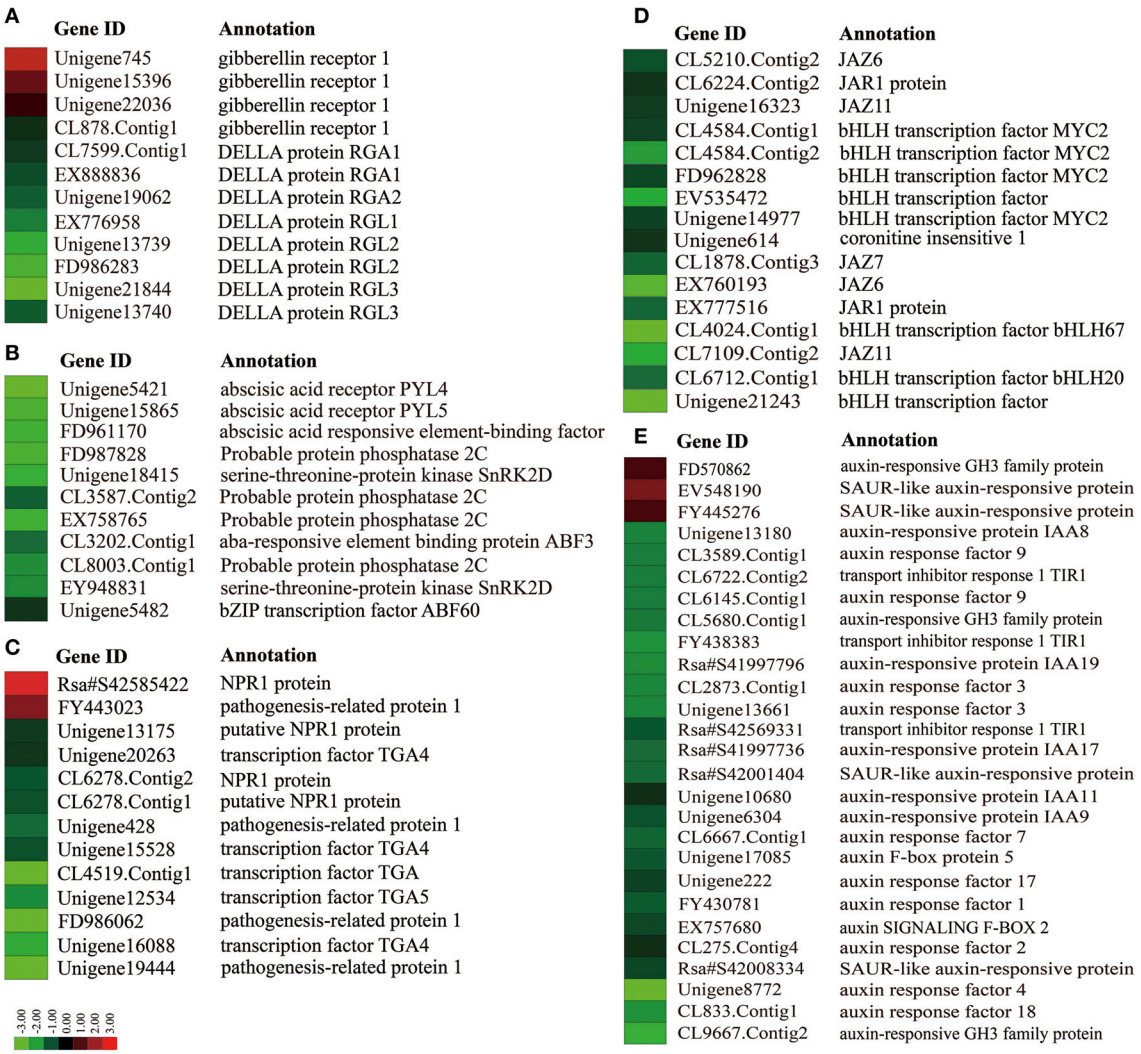


FIGURE 3 | Heat map diagram of expression patterns for DEGs involved in some phytohormone signaling pathways, including GA (A), ABA (B), SA (C), JA (D), and auxin (E). Red and green colors indicate up- and down-regulated genes in NAU-RS library as compared with NAU-VS library, respectively.

regulated by various endogenous and environmental signals integrating into a complexity of flowering regulation (Amasino and Michaels, 2010; Srikanth and Schmid, 2011). Recent advances in flowering genes and regulatory networks have greatly enhanced our knowledge of molecular basis underlying bolting and flowering-time control in Brassicaceae crops. However, no studies on comprehensive identification of DEGs related to radish bolting and flowering have been reported, and the regulatory mechanism of bolting and flowering-time control remains largely unexplored in radish. In this study, two cDNA libraries from leaves of radish advanced inbred line 'NAU-LU127' at vegetative and reproductive stages were constructed, respectively. A list of DEGs related to phytohormone signaling and transition from vegetative growth to bolting and flowering were identified and comprehensively profiled.

The Roles of Plant Hormone Signaling in Bolting and Flowering

Plant hormones are endogenously occurring compounds that regulate multiple aspects of plant growth and development including flowering time (Davis, 2009; Santner and Estelle, 2009). Various phytohormones have been implicated in the developmental transition of flowering (Davis, 2009; Domagalska et al., 2010). The pathways of several hormones including auxin, GA, ABA, SA, and JA signaling were significantly enriched by pathway-based analysis in our study (**Table 3**).

GA pathway is one of the genetic flowering pathways, which could interact with several pathways and is integrated into the flowering regulatory complexity (Srikanth and Schmid, 2011). The role of GA pathway in flowering time has been thoroughly investigated in *Arabidopsis* and several fruit trees (Wilkie et al., 2008; Mutasa-Göttgens and Hedden, 2009). Many

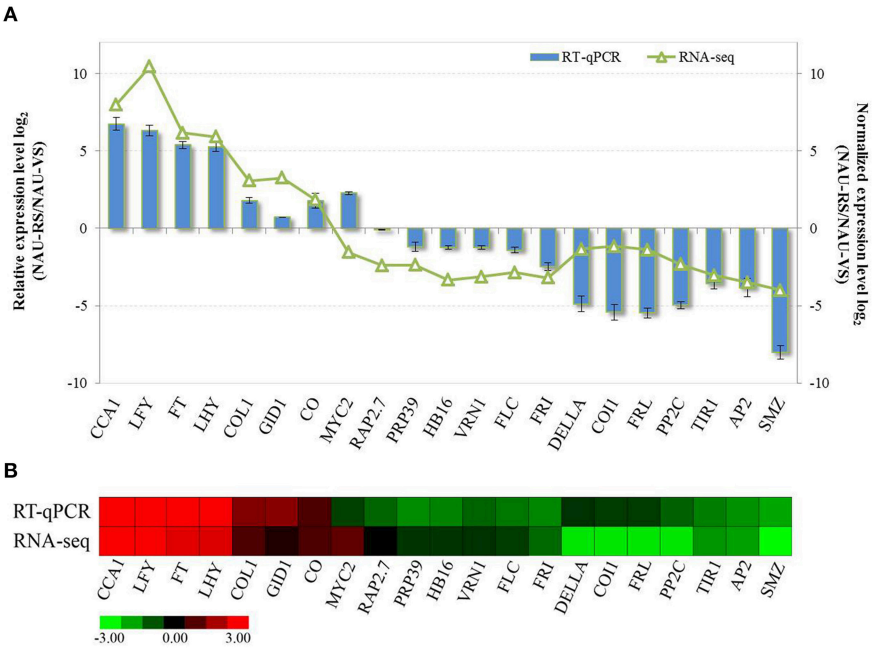


FIGURE 4 | The validation of expression levels of selected DEGs related to radish bolting and flowering. (A) The relative expression levels of selected DEGs were compared with the transcript abundances from DGE sequencing. **(B)** Heat map diagram of expression patterns of DEGs in radish. Red and green colors in indicate up- and down-regulated genes in NAU-RS library as compared with NAU-VS library, respectively.

genes related to GA metabolism and signaling were involved in GA-mediated regulatory process of flowering (Mutasa-Göttgens and Hedden, 2009; Domagalska et al., 2010). GA exerting its biological functions on floral transition and development is mainly dependent on the growth inhibitor DELLA proteins (Mutasa-Göttgens and Hedden, 2009). GA signaling promotes flowering through initiating the degradation of transcriptional regulator DELLA and activating the expression of *SOC1*, *AGL24* and *LFY* (Davis, 2009; Mutasa-Göttgens and Hedden, 2009). As expected, both the decreased transcript abundance and expression level of *DELLA* (CL7599.Contig1) were detected in radish reproductive stage compared with vegetative phase (**Figure 4**). The ABA pathway, which is antagonistic to GA, has been demonstrated to delay flowering through modulating DELLA activity and affecting the transcriptional expression of floral repressor *FLC* (Achard et al., 2006; Domagalska et al., 2010). In the current study, unique transcripts annotated as *PP2C* and *ABF*, ABA signaling components, were identified and differentially regulated during radish bolting and flowering (**Figure 3**), which is consistent with the results in litchi (Zhang et al., 2014c) and soybean (Wong et al., 2013). These findings suggested that the differential expressions of ABA signaling-related genes may be associated with the timing of radish transition to bolting and flowering.

Function of SA in accelerating transition to flowering is pronounced by SA-deficient mutants of *Arabidopsis* (Martínez et al., 2004). SA could negatively regulate the floral repressor *FLC* and activate the flowering promoter *FT* which strongly highlights the positive role of SA in flowering transition (Martínez et al., 2004). SA promotes the activation of NON-EXPRESSOR OF PR-1 (NPR1) proteins, whose interaction with TGA transcription factors could induce the expression of PR genes (Wu et al., 2012). Moreover, JA is also implicated in flowering regulatory process and delays flowering in *Arabidopsis* (Krajnčič et al., 2006; Riboni et al., 2014). JA signaling pathway has been involved in three molecular elements including JA receptor gene *COI1*, transcriptional repressor JAZ protein and some transcription factors, e.g., the bHLH family (Krajnčič et al., 2006). Notably, recent studies have demonstrated the regulatory role of *COI1* in delaying flowering mediating the repressed expression of *FT* (Zhai et al., 2015). In this study, some transcripts belonging to the main components of SA and JA signaling were found, including *NPR1*, *TGA*, *PR*, *JAZ*, *COI1*, and *MYC2* (**Table 3**). In addition, previous studies reveal that auxin is necessary for flower initiation and floral organ identity (Cheng and Zhao, 2007). We also found the critical genes related to auxin signaling such as *AUX1*, *SAUR*, *TIR1*, and *ARFs* (**Table 3**; **Table S4**). Overall, these results reveal that phytohormone-mediated transcriptional reprogramming are crucial to the transition of bolting and flowering and participate in its regulatory network of radish. The characterization of critical genes in plant hormone signaling pathways would greatly help to illuminate the complex genetic network of bolting and flowering in radish.

The Complex Bolting and Flowering Regulatory Network in Radish

Multiple genetic flowering pathways integrating endogenous and environmental signals determine the transition from vegetative

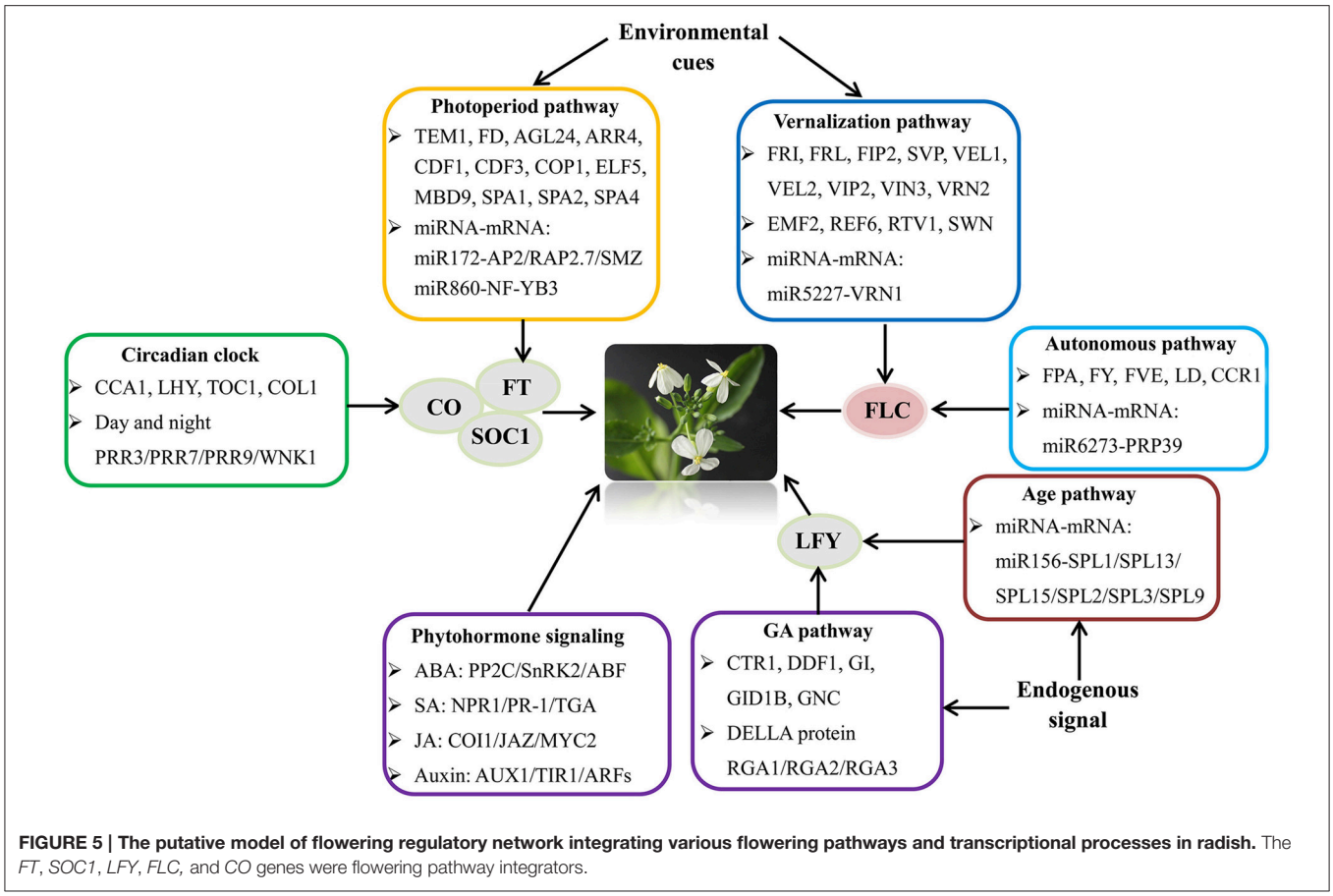


FIGURE 5 | The putative model of flowering regulatory network integrating various flowering pathways and transcriptional processes in radish. The *FT*, *SOC1*, *LFY*, *FLC,* and *CO* genes were flowering pathway integrators.

growth to reproductive development. Studies in *Arabidopsis* have revealed the participation of more than 200 flowering-related genes in the intricate regulatory network (Fornara et al., 2010; Srikanth and Schmid, 2011). In this study, 95 candidate genes related to bolting and flowering were isolated and involved in five major flowering pathways within genetic regulatory network (**Table S5**; **Figure 5**). It is inferred that known genetic pathways and critical flowering genes may conservatively present in radish, being consistent with the reports in maize (Dong et al., 2012), soybean (Jung et al., 2012), and citrus (Zhang et al., 2011). Gene expression profiling revealed that these genes were differentially expressed between NAU-VS and NAU-RS libraries, suggesting their putative important roles in radish bolting and flowering.

The complex regulatory network of *Arabidopsis* is composed of five major converging pathways (Fornara et al., 2010; Srikanth and Schmid, 2011). It is believed that endogenous developmental signals such as developmental stages of plants and phytohormones monitor flowering time through age, autonomous and GA pathways, while environmental cues regulate flowering time through the photoperiod and vernalization pathways in response to day length or temperature (Srikanth and Schmid, 2011; Capovilla et al., 2015). The signals from photoperiodic process are converted into the transcriptional regulation of key genes such as *FT*, *CO*, *AP1*, and *AP2* to affect flowering time (Kikuchi and Handa, 2009; Amasino, 2010; Srikanth and Schmid, 2011). The florigen gene *FT* as a floral integrator is central for the photoperiodic flowering pathway of long-day plant *Arabidopsis*, which is perceived in leaves and transported to the shoot apex initiating floral transition (Huang et al., 2005; Parcy, 2005). The role of *FT* in promoting flowering has been proven by mutants and overexpressed transgenic analysis in *Arabidopsis* (Amasino, 2010; Srikanth and Schmid, 2011). As expected, the homolog of *FT* (FD571044) was up-regulated in reproductive stage of radish (**Figure 4**), indicating that the *RsFT* gene could positively regulate the development transition of bolting and flowering (**Figure 5**). Under long-day condition, the *FT* expression is activated by *CO*, which is a floral activator and modulated by the circadian clock and day length (Suárez-López et al., 2001; Amasino, 2010; Johansson and Staiger, 2015). The link between circadian clock and flowering control may be mainly mediated by the transcriptional expression of *CO* (Fujiwara et al., 2008; Johansson and Staiger, 2015). In *Arabidopsis*, two essential circadian clock components *LATE ELONGATED HYPOCOTYL* (*LHY*) and *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) function in photoperiodic flowering and regulate flowering pathway by controlling the rhythmic expression of *CO* and *FT* (Fujiwara et al., 2008). In this study, some transcripts belonging to *CO*, *CCA1* and *LHY* homologs were found to be up-expressed in reproductive phase with DGE sequencing and RT-qPCR analysis (**Figure 4**), suggesting the critical roles of these genes in the transition of radish bolting and flowering.

The vernalization and autonomous pathways converge on the flowering repressor *FLC*, and many genes involved in these two pathways could control flowering time through affecting *FLC* expression (Amasino, 2010). The high level of *FLC* delays flowering and requires its activator *FRI* (Michaels and Amasino, 1999; Choi et al., 2011). Recently, several naturally occurring spliced transcripts of *FLC* were found and isolated from *B. rapa* (Yuan et al., 2009) and orange (Zhang et al., 2009), which were proven to be associated with variations in flowering time. The transcriptional co-expression analysis in *B. rapa* indicated that *BrFLC2* may be the major regulator of flowering time in genetic flowering network (Xiao et al., 2013). In this study, we found putative homologs of *FLC* (CL1584.Contig3) from late-bolting radish, which was down-regulated in reproductive stage compared with vegetative stage, with similar patterns being detected in *FRI* and *FRL* (**Figure 4**). In addition, similar results were found in other homologous genes in vernalization pathway, including *FIP2*, *EMF2*, *VRN1,* and *VRN2* (**Table S5**). These results indicate that the genetic elements of the vernalization pathway may be of importance for the manipulation of radish bolting and flowering time.

Furthermore, miRNAs as central regulators of gene expression have been shown to be implicated in multiple genetic pathways governing flowering time (Spanudakis and Jackson, 2014; Wang, 2014). The newly defined age pathway of flowering, which is controlled by miR156 and its target *SPL* transcription factors (Wang et al., 2009b), regulates flowering time and interacts with vernalization, photoperiodic and GA pathways (Zhou et al., 2013; Spanudakis and Jackson, 2014; Wang, 2014). Several members of *SPL* family were identified in this study, including *SPL1*, *SPL2*, *SPL3*, *SPL9*, *SPL13,* and *SPL15* (**Table S5**). It was known that miR172 is down-regulated by the age-dependent expression of *SPL9* (Wu et al., 2009; Spanudakis and Jackson, 2014). The target genes of miR172 are a class of *AP2*-like transcription factors including *AP2*, *TARGET OF EAT 1-3* (*TOE1-3*), *SCHLAFMÜTZE* (*SMZ*), and *SCHNARCHZAPFEN* (*SNZ*), which act as floral repressors (Zhu and Helliwell, 2010). The levels of these *AP2*-like genes are relatively high during plant seedling stage and decline with plant development, ultimately relieving the repression of flowering to trigger flowering (Aukerman and Sakai, 2003; Zhu and Helliwell, 2010). Consistent with these evidences, the down-expressed patterns of *AP2* (CL1275.Contig1), *SMZ* (Rsa#S42015352), and *RAP2.7* (CL2600.Contig3) were detected at reproductive stage in this study (**Figure 4**). Moreover, correlation analysis revealed that some bolting and flowering-related DEGs were targeted by specific miRNAs forming the transcriptional model of miRNA-mRNA pairs (**Table S6**). These findings reveal that some miRNA-DEG models including miR5227-*VRN1*, miR6273-*PRP39,* and miR860-*NF-YB3* are crucial participators and integrated into the intricate genetic networks of bolting and flowering in radish (**Figure 5**).

CONCLUSIONS

In summary, RNA-Seq technology was employed to systematically identify DEGs at transcriptome-wide level during radish transition from vegetative growth to bolting and flowering in this study. To our knowledge, this is the first investigation to illustrate the expression profiles of bolting-related genes and dissect the bolting and flowering regulatory network in radish. In this study, a total of 5922 DEGs were identified from late-bolting radish leaves. Several candidate genes related to plant hormone signal and bolting and flowering regulatory pathways were characterized and implicated in the complex networks of bolting and flowering regulation. Correlation analysis suggested that the miRNA-mRNA regulatory models played pivotal roles in determining bolting and flowering time. Moreover, a schematic regulatory network of radish bolting and flowering was put forward for characterization of DEGs and miRNAs. These results provided essential information for genetic control of radish bolting and flowering, and would facilitate unraveling the molecular regulatory mechanism underlying bolting and flowering in radish and other root vegetable crops.

AUTHOR CONTRIBUTIONS

SN, CL, and LL designed the research. SN, XS, and MT conducted experiments. SN, LX, and YW participated in the design of the study and performed the statistical analysis. SN analyzed data and wrote the manuscript. LL and EM helped with the revision of manuscript. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

This work was partially supported by grants from the National Natural Science Foundation of China (31171956, 31372064), the National Key Technologies R & D Program of China (2012BAD02B01) and Key Technologies R & D Program of Jiangsu Province (BE2013429).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00682

Table S1 | All the identified DEGs in NAU-VS and NAU-RS libraries.

Table S2 | GO enrichment analysis for differentially expressed transcripts with corrected $P \leq 0.05$.

Table S3 | KEGG pathway analysis of differentially expressed transcripts in radish.

Table S4 | The DEGs involved in plant hormone signal transduction pathway.

Table S5 | The DEGs involved in the transition of vegetative growth to bolting in radish.

Table S6 | The identified DEG and miRNA pairs during radish bolting and flowering.

Table S7 | The primers of DEGs for RT-qPCR in radish.

Figure S1 | The identified genes involved in plant hormone signal transduction by KEGG analysis.

REFERENCES

Achard, P., Cheng, H., Grauwe, L., Decat, J., Schoutteten, H., Moritz, T., et al. (2006). Integration of plant responses to environmentally activated phytohormonal signals. *Science* 311, 91–94. doi: 10.1126/science.1118642

Amasino, R. (2010). Seasonal and developmental timing of flowering. *Plant J.* 61, 1001–1013. doi: 10.1111/j.1365-313X.2010.04148.x

Amasino, R. M., and Michaels, S. D. (2010). The timing of flowering. *Plant Physiol.* 154, 516–520. doi: 10.1104/pp.110.161653

Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7, 986–995. doi: 10.1101/gr.7.10.986

Aukerman, M. J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2-like* target genes. *Plant Cell* 15, 2730–2741. doi: 10.1105/tpc.016238

Bai, S., Saito, T., Sakamoto, D., Ito, A., Fujii, H., and Moriguchi, T. (2013). Transcriptome analysis of japanese pear (*Pyrus pyrifolia* Nakai) flower buds transitioning through endodormancy. *Plant Cell Physiol.* 54, 1132–1151. doi: 10.1093/pcp/pct067

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behav. Brain. Res.* 125, 279–284. doi: 10.1016/S0166-4328(01)00297-2

Capovilla, G., Schmid, M., and Posé, D. (2015). Control of flowering by ambient temperature. *J. Exp. Bot.* 66, 59–69. doi: 10.1093/jxb/eru416

Cheng, Y., and Zhao, Y. (2007). A role for auxin in flower development. *J. Integr. Plant Biol.* 49, 99–104. doi: 10.1111/j.1744-7909.2006.00412.x

Choi, K., Kim, J., Hwang, H., Kim, S., Park, C., Kim, S., et al. (2011). The FRIGIDA complex activates transcription of *FLC*, a strong flowering repressor in *Arabidopsis*, by recruiting chromatin modification factors. *Plant Cell* 23, 289–303. doi: 10.1105/tpc.110.075911

Conesa, A., Götz, S., García-Gómez, J., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Davis, S. J. (2009). Integrating hormones into the floral-transition pathway of *Arabidopsis thaliana*. *Plant Cell Environ.* 32, 1201–1210. doi: 10.1111/j.1365-3040.2009.01968.x

Domagalska, M. A., Sarnowska, E., Nagy, F., and Davis, S. J. (2010). Genetic analyses of interactions among gibberellin, abscisic acid, and brassinosteroids in the control of flowering time in *Arabidopsis thaliana*. *PLoS ONE* 5:e14012. doi: 10.1371/journal.pone.0014012

Dong, Z., Danilevskaya, O., Abadie, T., Messina, C., Coles, N., and Cooper, M. (2012). A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. *PLoS ONE* 7:e43450. doi: 10.1371/journal.pone.0043450

Fornara, F., de Montaigu, A., and Coupland, G. (2010). SnapShot: control of flowering in *Arabidopsis*. *Cell* 141, 550. doi: 10.1016/j.cell.2010.04.024

Fujiwara, S., Oda, A., Yoshida, R., Niinuma, K., Miyata, K., Tomozoe, Y., et al. (2008). Circadian clock proteins LHY and CCA1 regulate SVP protein accumulation to control flowering in *Arabidopsis*. *Plant Cell* 20, 2960–2971. doi: 10.1105/tpc.108.061531

Gao, J., Zhang, Y., Zhang, C., Qi, F., Li, X., Mu, S., et al. (2014). Characterization of the floral transcriptome of moso bamboo (*Phyllostachys edulis*) at different flowering developmental stages by transcriptome sequencing and RNA-Seq analysis. *PLoS ONE* 9:e98910. doi: 10.1371/journal.pone.0098910

Huang, T., Böhlenius, H., Eriksson, S., Parcy, F., and Nilsson, O. (2005). The mRNA of the *Arabidopsis* gene *FT* moves from leaf to shoot apex and induces flowering. *Science* 309, 1694–1696. doi: 10.1126/science.1117768

Johansson, M., and Staiger, D. (2015). Time to flower: interplay between photoperiod and the circadian clock. *J. Exp. Bot.* 66, 719–730. doi: 10.1093/jxb/eru441

Jung, C. H., Wong, C. E., Singh, M. B., and Bhalla, P. L. (2012). Comparative genomic analysis of soybean flowering genes. *PLoS ONE* 7:e38250. doi: 10.1371/journal.pone.0038250

Kazan, K., and Lyons, R. (2015). The link between flowering time and stress tolerance. *J. Exp. Bot.* 67, 295–301. doi: 10.1093/jxb/erv441

Kikuchi, R., and Handa, H. (2009). Photoperiodic control of flowering in barley. *Breeding Sci.* 59, 546–552. doi: 10.1270/jsbbs.59.546

Krajnčič, B., Kristl, J., and Janžekovič, I. (2006). Possible role of jasmonic acid in the regulation of floral induction, evocation and floral differentiation in *Lemna minor* L. *Plant Physiol. Bioch.* 44, 752–758. doi: 10.1016/j.plaphy.2006.10.029

Kumimoto, R. W., Adam, L., Hymus, G. J., Repetti, P. P., Reuber, T. L., Marion, C. M., et al. (2008). The Nuclear Factor Y subunits NF-YB2 and NF-YB3 play additive roles in the promotion of flowering by inductive long-day photoperiods in *Arabidopsis*. *Planta* 228, 709–723. doi: 10.1007/s00425-008-0773-6

Lister, R., Gregory, B. D., and Ecker, J. R. (2009). Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr. Opin. Plant Biol.* 12, 107–118. doi: 10.1016/j.pbi.2008.11.004

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Martínez, C., Pons, E., Prats, G., and León, J. (2004). Salicylic acid regulates flowering time and links defence responses and reproductive development. *Plant J.* 37, 209–217. doi: 10.1046/j.1365-313X.2003.01954.x

Michaels, S., and Amasino, R. (1999). *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11, 949–956. doi: 10.1105/tpc.11.5.949

Moon, J., Lee, H., Kim, M., and Lee, I. (2005). Analysis of flowering pathway integrators in *Arabidopsis*. *Plant Cell Physiol.* 46, 292–299. doi: 10.1093/pcp/pci024

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Mutasa-Göttgens, E., and Hedden, P. (2009). Gibberellin as a factor in floral regulatory networks. *J. Exp. Bot.* 60, 1979–1989. doi: 10.1093/jxb/erp040

Nie, S., Xu, L., Wang, Y., Huang, D., Muleke, E. M., Sun, X., et al. (2015). Identification of bolting-related microRNAs and their targets reveals complex miRNA-mediated flowering-time regulatory networks in radish (*Raphanus sativus* L.). *Sci. Rep.* 5:14034. doi: 10.1038/srep14034

Nie, S., Li, C., Xu, L., Wang, Y., Huang, D., Muleke, E. M., et al. (in press). *De novo* transcriptome analysis and identification of critical genes involved in bolting and flowering in radish (*Raphanus sativus* L.). *BMC Genomics* 17. doi: 10.1186/s12864-016-2633-2

Oshlack, A., Robinson, M. D., and Young, M. D. (2010). From RNA-Seq reads to differential expression results. *Genome Biol.* 11, 220. doi: 10.1186/gb-2010-11-12-220

Parcy, M. (2005). Flowering: a time for integration. *Int. J. Dev. Biol.* 49, 585–593. doi: 10.1387/ijdb.041930fp

Riboni, M., Robustelli, A., Galbiati, M., Tonelli, C., and Conti, L. (2014). Environmental stress and flowering time: the photoperiodic connection. *Plant Signal. Behav.* 9:e29036. doi: 10.4161/psb.29036

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Robinson, M. D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-Seq data. *Genome Biol.* 11, 79–82. doi: 10.1186/gb-2010-11-3-r25

Saldanha, A. J. (2004). Java Treeview-extensible visualization of microarray data. *Bioinformatics* 20, 3246–3248. doi: 10.1093/bioinformatics/bth349

Santner, A., and Estelle, M. (2009). Recent advances and emerging trends in plant hormone signalling. *Nature* 459, 1071–1078. doi: 10.1038/nature08122

Spanudakis, E., and Jackson, S. (2014). The role of microRNAs in the control of flowering time. *J. Exp. Bot.* 65, 365–380. doi: 10.1093/jxb/ert453

Srikanth, A., and Schmid, M. (2011). Regulation of flowering time: all roads lead to Rome. *Cell Mol. Life Sci.* 68, 2013–2037. doi: 10.1007/s00018-011-0673-y

Suárez-López, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., and Coupland, G. (2001). CONSTANS mediates between the circadian clock and the control of flowering in *Arabidopsis*. *Nature* 410, 1116–1120. doi: 10.1038/35074138

Tao, X., Gu, Y., Jiang, Y., Zhang, Y., and Wang, H. (2013). Transcriptome analysis to identify putative floral-specific genes and flowering regulatory-related genes of sweet potato. *Biosci. Biotech. Bioch.* 77, 2169–2174. doi: 10.1271/bbb.130218

Wang, C., Tian, Q., Hou, Z., Mucha, M., Aukerman, M., and Olsen, O. (2007). The *Arabidopsis thaliana ATPRP39-1* gene, encoding a tetratricopeptide repeat protein with similarity to the yeast pre-mRNA processing protein PRP39,

affects flowering time. *Plant Cell Rep.* 26, 1357–1366. doi: 10.1007/s00299-007-0336-5

Wang, L., Feng, Z., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-Seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612

Wang, Z., Gerstein, M., and Snyder, M. (2009a). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. doi: 10.1038/nrg2484

Wang, J. W., Czech, B., and Weigel, D. (2009b). miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell* 138, 738–749. doi: 10.1016/j.cell.2009.06.014

Wang, J. W. (2014). Regulation of flowering time by the miR156-mediated age pathway. *J. Exp. Bot.* 65, 4723–4730. doi: 10.1093/jxb/eru246

Wang, Y., Pan, Y., Liu, Z., Zhu, X., Zhai, L., Xu, L., et al. (2013). *De novo* transcriptome sequencing of radish (*Raphanus sativus* L.) and analysis of major genes involved in glucosinolate metabolism. *BMC Genomics* 14:836. doi: 10.1186/1471-2164-14-836

Wilkie, J. D., Sedgley, M., and Olesen, T. (2008). Regulation of floral initiation in horticultural trees. *J. Exp. Bot.* 59, 3215–3228. doi: 10.1093/jxb/ern188

Wong, C. E., Singh, M. B., and Bhalla, P. L. (2013). The dynamics of soybean leaf and shoot apical meristem transcriptome undergoing floral initiation process. *PLoS ONE* 8:65319. doi: 10.1371/journal.pone.0065319

Wu, G., Park, M. Y., Conway, S. R., Wang, J., Weigel, D., and Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138, 750–759. doi: 10.1016/j.cell.2009.06.031

Wu, Y., Zhang, D., Chu, J. Y., Boyle, P., Wang, Y., Brindle, I. D., et al. (2012). The *Arabidopsis* NPR1 protein is a receptor for the plant defense hormone salicylic acid. *Cell Rep.* 1, 639–647. doi: 10.1016/j.celrep.2012.05.008

Xiao, D., Zhao, J., Hou, X., Basnet, R. K., Carpio, D., Zhang, N., et al. (2013). The *Brassica rapa FLC* homologue *FLC2* is a key regulator of flowering time, identified through transcriptional co-expression networks. *J. Exp. Bot.* 64, 4503–4516. doi: 10.1093/jxb/ert264

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. doi: 10.1093/nar/gkr483

Xu, L., Wang, Y., Liu, W., Wang, J., Zhu, X., Zhang, K., et al. (2015). *De novo* sequencing of root transcriptome reveals complex cadmium-responsive regulatory networks in radish (*Raphanus sativus* L.). *Plant Sci.* 236, 313–323. doi: 10.1016/j.plantsci.2015.04.015

Xu, Y., Zhu, X., Gong, Y., Xu, L., Wang, Y., and Liu, L. (2012). Evaluation of reference genes for gene expression studies in radish (*Raphanus sativus* L.) using quantitative real-time PCR. *Biochem. Biophys. Res. Commun.* 424, 398–403. doi: 10.1016/j.bbrc.2012.06.119

Yaish, M. W., Colasanti, J., and Rothstein, S. J. (2011). The role of epigenetic processes in controlling flowering time in plants exposed to stress. *J. Exp. Bot.* 62, 3727–3735. doi: 10.1093/jxb/err177

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, W293–W297. doi: 10.1093/nar/gkl031

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-Seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

Yuan, Y., Wu, J., Sun, R., Zhang, X., Xu, D., Bonnema, G., et al. (2009). A naturally occurring splicing site mutation in the *Brassica rapa FLC1* gene is associated with variation in flowering time. *J. Exp. Bot.* 60, 1299–1308. doi: 10.1093/jxb/erp010

Zhai, Q., Zhang, X., Wu, F., Feng, H., Deng, L., Xu, L., et al. (2015). Transcriptional mechanism of Jasmonate receptor COI1-mediated delay of flowering time in *Arabidopsis*. *Plant Cell* 27, 2814–2828. doi: 10.1105/tpc.15.00619

Zhang, J., Li, Z., Mei, L., Yao, J., and Hu, C. (2009). *PtFLC* homolog from trifoliate orange (*Poncirus trifoliata*) is regulated by alternative splicing and experiences seasonal fluctuation in expression level. *Planta* 229, 847–859. doi: 10.1007/s00425-008-0885-z

Zhang, J., Ai, X., Sun, L., Zhang, D., Guo, W., Deng, X., et al. (2011). Transcriptome profile analysis of flowering molecular processes of early flowering trifoliate orange mutant and the wild-type [*Poncirus trifoliate* (L.) Raf.] by massively parallel signature sequencing. *BMC Genomics* 12:63. doi: 10.1186/1471-2164-12-63

Zhang, Y., Peng, L., Wu, Y., Shen, Y., Wu, X., and Wang, J. (2014a). Analysis of global gene expression profiles to identify differentially expressed genes critical for embryo development in *Brassica rapa*. *Plant Mol. Biol.* 86, 425–442. doi: 10.1007/s11103-014-0238-1

Zhang, Z., Wang, P., Li, Y., Ma, L., Li, L., Yang, R., et al. (2014b). Global transcriptome analysis and identification of the flowering regulatory genes expressed in leaves of *Lagerstroemia indica*. *DNA and Cell Biol.* 33, 680–688. doi: 10.1089/dna.2014.2469

Zhang, H., Wei, Y., Shen, J., Lai, B., Huang, X., Ding, F., et al. (2014c). Transcriptomic analysis of floral initiation in litchi (*Litchi chinensis* Sonn.) based on *de novo* RNA sequencing. *Plant Cell Rep.* 33, 1723–1735. doi: 10.1007/s00299-014-1650-3

Zhang, L., Jia, H., Yin, Y., Wu, G., Xia, H., Wang, X., et al. (2013). Transcriptome analysis of leaf tissue of *Raphanus sativus* by RNA sequencing. *PLoS ONE* 8:e80350. doi: 10.1371/journal.pone.0080350

Zhou, C., Zhang, T., Wang, X., Yu, S., Lian, H., Tang, H., et al. (2013). Molecular basis of age-dependent vernalization in *Cardamine flexuosa*. *Science* 340, 1097–1100. doi: 10.1126/science.1234340

Zhu, Q. H., and Helliwell, C. A. (2010). Regulation of flowering time and floral patterning by miR172. *J. Exp. Bot.* 62, 487–495. doi: 10.1093/jxb/erq295

Zhu, Y., Li, Y., Xin, D., Chen, W., Shao, X., Wang, Y., et al. (2015). RNA-Seq-based transcriptome analysis of dormant flower buds of Chinese cherry (*Prunus pseudocerasus*). *Gene* 555, 362–376. doi: 10.1016/j.gene.2014.11.032

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Nie, Li, Wang, Xu, Muleke, Tang, Sun and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 May 2016
doi: 10.3389/fpls.2016.00585

De novo Taproot Transcriptome Sequencing and Analysis of Major Genes Involved in Sucrose Metabolism in Radish (*Raphanus sativus* L.)

*Rugang Yu[1,2], Liang Xu[1], Wei Zhang[1], Yan Wang[1], Xiaobo Luo[1], Ronghua Wang[1], Xianwen Zhu[3], Yang Xie[1], Benard Karanja[1] and Liwang Liu[1]**

[1] National Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing, China, [2] School of Life Science, Huaibei Normal University, Huaibei, China, [3] Department of Plant Sciences, North Dakota State University, Fargo, ND, USA

OPEN ACCESS

Edited by:
Narendra Tuteja,
International Centre for Genetic Engineering and Biotechnology, India

Reviewed by:
Maoteng Li,
Huazhong University of Science and Technology, China
Maria D. Logacheva,
Lomonosov Moscow State University, Russia

***Correspondence:**
Liwang Liu
nauliulw@njau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: 10 July 2015
Accepted: 15 April 2016
Published: 17 May 2016

Citation:
Yu R, Xu L, Zhang W, Wang Y, Luo X, Wang R, Zhu X, Xie Y, Karanja B and Liu L (2016) De novo Taproot Transcriptome Sequencing and Analysis of Major Genes Involved in Sucrose Metabolism in Radish (Raphanus sativus L.). Front. Plant Sci. 7:585. doi: 10.3389/fpls.2016.00585

Radish (*Raphanus sativus* L.) is an important annual or biennial root vegetable crop. The fleshy taproot comprises the main edible portion of the plant with high nutrition and medical value. Molecular biology study of radish begun rather later, and lacks sufficient transcriptomic and genomic data in pubic databases for understanding of the molecular mechanism during the radish taproot formation. To develop a comprehensive overview of the 'NAU-YH' root transcriptome, a cDNA library, prepared from three equally mixed RNA of taproots at different developmental stages including pre-cortex splitting stage, cortex splitting stage, and expanding stage was sequenced using high-throughput Illumina RNA sequencing. From approximately 51 million clean reads, a total of 70,168 unigenes with a total length of 50.28 Mb, an average length of 717 bp and a N50 of 994 bp were obtained. In total, 63,991 (about 91.20% of the assembled unigenes) unigenes were successfully annotated in five public databases including NR, GO, COG, KEGG, and Nt. GO analysis revealed that the majority of these unigenes were predominately involved in basic physiological and metabolic processes, catalytic, binding, and cellular process. In addition, a total of 103 unigenes encoding eight enzymes involved in the sucrose metabolism related pathways were also identified by KEGG pathway analysis. Sucrose synthase (29 unigenes), invertase (17 unigenes), sucrose-phosphate synthase (16 unigenes), fructokinase (17 unigenes), and hexokinase (11 unigenes) ranked top five in these eight key enzymes. From which, two genes (*RsSuSy1*, *RsSPS1*) were validated by T-A cloning and sequenced, while the expression of six unigenes were profiled with RT-qPCR analysis. These results would be served as an important public reference platform to identify the related key genes during taproot thickening and facilitate the dissection of molecular mechanisms underlying taproot formation in radish.

Keywords: *Raphanus sativus*, taproot, RNA-seq, transcriptome, sucrose metabolism

INTRODUCTION

Radish (*Raphanus sativus* L., $2n = 2x = 18$) is an important root vegetable crop belonging to the Brassicaceae family grown all over the world, especially in East Asia (Johnston et al., 2005; Wang and He, 2005). The fleshy taproot is a key organ for the direct yield and quality of radish, and its formation and development is a complex biological processes involving morphogenesis and dry matter accumulation (Wang and He, 2005). During this process, an abundance of storage compounds are synthesized, including carbohydrates, ascorbic acid, folic acid, potassium, vitamin B6, riboflavin, magnesium and sulforaphane, which mainly could determine the economic value of radish taproot and provide nutrients and medicinal function for human beings (Curtis, 2003; Gutiérrez and Perez, 2004; Chaturvedi, 2008; Wang et al., 2013). Hence, understanding the processes regulating the root formation and development is of particular importance.

To date, large amounts of transcriptomic and genomic sequences have been provided in model plants, such as Arabidopsis, Antirrhinum and rice, which have greatly helped the understanding of the complexity of growth and development in higher plants. For radish, many researches reported that the genomic information have recently been analyzed. In recent years, the genomic and transcriptome information of radish have been extensively clarified. For instance, two leaf and two root transcriptomes from radish were reported (Wang et al., 2012, 2013; Zhang et al., 2013; Wu et al., 2015), and some critical genes associated with glucosinolate metabolism and heavy metal stress response were identified (Wang et al., 2013). In addition, 314,823 expressed sequence tags (ESTs), 31,935 nucleotide sequences and 16 genome survey sequences (GSS) were stored in NCBI for radish (http://www.ncbi.nlm.nih.gov/nucest/?term=radish) (March 7th, 2014). More recently, the draft genome sequences of *R. sativus* have been assembled and published (Kitashiba et al., 2014). These data might provide the useful database for genomic and functional investigation on some important horticultural traits in radish. However, the formation and development of taproot is a complex biological process in radish. The radish advanced inbred line, 'NAU-YH' with a taproot in very small size (maximum diameter <3.0 cm at maturity), is a very suitable genotype for taproot development investigation. However, the genome and the transcriptome of 'NAU-YH' were not sequenced, and the related genomic information was still unavailable. The resulting data of transcriptome sequencing of this genotype would be useful for further molecular investigation on taproot development.

Sucrose is the major product of photosynthesis (Ruan, 2012). Generally, it is the main form of assimilated carbon to be transported from "source" to "sink" organs in higher plants (Farrar et al., 2000). In addition, sucrose is not only the source of carbon skeletons which may involve in the synthesis essential metabolite compounds including starch, cellulose and proteins (Weber et al., 1997; Cheng and Chourey, 1999; Babb and Haigler, 2001), but also an important signal molecule in plants that regulates the expression of microRNA (Yang et al., 2013), transcription factors (Xiong et al., 2013), plant hormone (Stokes et al., 2013), and other genes (Ruan, 2014). Therefore, sucrose metabolism plays important roles in plant growth and development.

There are three key enzymes responsible for sucrose synthesis and degradation in plants, including invertase (INV, EC 3.2.1.26, involved in sucrose degradation), sucrose synthase (SuSy, EC 2.4.1.13, involved in sucrose degradation), and sucrose-phosphate synthase (SPS, EC 2.3.1.14, involved in sucrose synthesis; Ren and Zhang, 2013). During last decade, extensive knowledge of sucrose metabolism has been studied by cloning and characterizing the genes encoding key enzymes in various plant species. For example, *SuSy* gene activity was found to be related to sink energy in tomato fruit (Wang et al., 1993); The gene structure, expression and regulation, and the physiological functions of the key enzymes involved in sucrose metabolism in maize were reviewed by Ren and Zhang (2013); Li and Zhang (2003) reported that SuSy was the most actively expressed enzyme in sucrose metabolism in developing storage root and was correlated with sink strength, while invertase was active at cell formation stages in Sweet Potato. The fleshy taproot of radish is one of major sink organ. Its growth and development requires an increase in sink activity, which is obtained by activating sucrose metabolism. Usuda et al. (1999) found that sink activity was strongly related to the level and activity of sucrose synthase but not to the activity of invertase. Wang et al. (2007) also found that the activities of SuSy were similar to sink activities in all lines. These results suggested that these enzymes might be associated with the developmental of the sink organ of radish. To date, although several studies have reported the role of sucrose metabolism in radish taproot thickening growth (Rouhier and Usuda, 2001; Wang et al., 2007), molecular mechanisms underlying sucrose metabolism remains unclear, especially for identification and evaluation of the full range of gene involved in sucrose metabolism in radish taproot.

Next-generation sequencing (NGS)-based RNA sequencing for transcriptome methods (RNA-seq) has been proven to be an effective method to analyze functional gene variation, and dramatically improve the speed and efficiency of gene discovery. (Angeloni et al., 2011; Hyun et al., 2012; Ward et al., 2012). The aim of this study was to obtain a comprehensive survey of transcripts associated with radish taproot formation. We utilize Illumina paired-end Solexa sequencing to conduct the *de novo* assembly and annotation of the 'NAU-YH' taproot transcriptome. According to KEGG pathway information, we first identified candidate genes of the key enzymes involved in sucrose metabolism and estimated the expression levels of these genes in different stages of taproot thickening. These results would provide important information for identifying the related key genes during taproot formation and facilitate further understanding of molecular mechanisms underlying taproot thickening in radish.

MATERIALS AND METHODS

Plant Material and RNA Extraction

The radish (*R. sativus* L.) advanced inbred line 'NAU-YH' was chosen for this study. Seeds were selected and germinated on

moist filter paper in darkness for 3 days. Then, seedlings were transplanted into plastic pots containing 1:1 mixture of soil and peat substrate, and cultured in the greenhouse at Nanjing Agricultural University. The development of cortex splitting is an important signal of the initiation of thickening growth of taproot in radish due to the cortex cells cannot divide and expand (Wang and He, 2005). Moreover, according to the 'NAU-YH' radish established morphological traits, the root cortex split initiated about 12 days after sowing (DAS), and the full root cortex splitting was achieved over a period of 22 DAS. The growth of root indicated rapidly thickening in the 22 to 42 DAS, then continued into a slowly thickening period. Therefore, samples of taproots were collected at three different development stages: 10 (DAS) (Stage 1, pre-cortex splitting stage), 20 DAS (Stage 2, cortex splitting stage), and 40 DAS (Stage 3, expanding stage) in this study. The subsamples of taproot, stem and leaf tissues were collected at 10, 20, 40, and 50 DAS, respectively for RT-qPCR verification. All samples were frozen in liquid nitrogen and stored at –80°C for further use.

Total RNA was extracted separately from the three taproot samples using Trizol regent (Invitrogen, USA) following the manufacturer's protocol. After the RNase-free DNase I (Takara, Japan) treatment, for cDNA preparation, a total 20 μg of RNA was mixed equally from each of the three taproot samples.

cDNA Library Construction and Sequencing

After the total RNA extraction, mRNA was purified from the 20 μg of RNA using Sera-mag Magnetic Oligo (dT) Beads (Thermo Fisher Scientific, USA). Then the purified mRNA was broken into small pieces using fragmentation buffer under elevated temperature. These short fragments as templates were used to synthesize first strand cDNA. Subsequently, the second-strand cDNA was synthesized using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, USA). The short cDNA fragments were purified with Qia-Quick PCR extraction kit and end-repair with EB buffer and ligation of A-tailing. Next, suitable fragments were selected as templates for PCR amplification to create the final cDNA library. Finally, after validating on an Agilent Technologie 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System, the cDNA library was sequenced at the Beijing Genomics Institute (BGI, Shenzhen, China) using Illumina HiSeq^{TM} 2000 sequencing platform. Image data outputs from sequencing machine were transformed by base calling into sequence data, which is called raw reads.

Data Filtering and *De novo* Assembly

The clean reads were generated by removing adaptor reads, empty reads, and low quality reads from the raw reads. Then, the clean reads were assembled using a *de novo* assembly program Trinity (Grabherr et al., 2011) with default K-mers = 25. Briefly, the process was done as previously described procedure (Wang et al., 2013). The clean reads with a certain length of overlap were firstly used to produce contigs. The reads were then mapped back to the contigs, and the paired-end reads was used to detect contigs from the same transcript as well as the distances between these contigs. To reduce any sequence redundancy, the contigs were further connected using Trinity after the paired-end reads, and sequences that could not be extended on either end were defined as unigenes. Finally, the unigenes were divided into two classes by gene family clustering. One is clusters, several unigenes with over 70% similarity between them, and the other unigenes were singletons.

Functional Annotation and Classification

The assembled unigene sequences were aligned by BLASTx to the publicly available protein databases which included NCBI non redundant protein (Nr), Gene Ontology (GO), Clusters of Orthologous Groups (COG), Swiss-Prot protein and the Kyoto Encyclopedia of Genes and Genomes (KEGG), and aligned by BLASTn to nucleotide databases (Nt) with an $E \leq 10^{-5}$. The best alignments in blast results were taken to decide the coding region sequences of the assembled unigenes. If the results from different databases conflicted with each other, a priority order of Nr, Swiss-Prot, KEGG and COG was followed. Meanwhile, if the assembled unigene sequences could not be aligned to any database, the software ESTScan was used to predict the protein coding sequence (CDS) and its sequence orientation (Iseli et al., 1999). And then, GO annotation of the unigenes was performed based on the best hits from Nr annotation using BLAST2GO program (Conesa et al., 2005), and the results of GO annotation were further used to conduct GO functional classification by WEGO software (Ye et al., 2006).

Gene Validation by T-A Cloning and Sequencing

According to the conserved region of radish EST sequences from radish cDNA library, the specific PCR primers of the two selected genes were designed to isolate sucrose metabolism related genes using Primer 5.0 software (Table S1). PCR was performed according to the method described previously (Wang et al., 2013). The PCR products were separated and ligated into the pMD18-T vector (Takara Bio Inc., China), and then transformed into *E. coli* DH5α. Positive clones were sequenced with ABI 3730 (Applied Bio systems, USA).

RT-qPCR Analysis

Six selected unigenes with crucial roles in sucrose metabolism were selected for RT-qPCR analysis using the SYBR Green Master ROX (Roche, Japan). The unigenes specific primers were designed using Beacon Designer 7.0 software (Table S1). Total RNAs were respectively extracted from taproot, stems and leaves in four different taproot development stages (10, 20, 40, and 50 DAS) using Trizol® Reagent (Invitrogen, USA) and then treated with PrimeScript® RT reagent Kit (Takara, Dalian, China) to reverse transcribe into cDNA. The amplification reactions were run in iCycler iQ real-time PCR detection system (BIO-RAD) according to previous reports (Xu et al., 2012). All reactions were performed in three replicates and the equation ratio = $2^{-\Delta\Delta C_T}$ was applied to calculate the relative expression level of the selected unigenes using *Actin* gene as the internal control gene. The data were analyzed using the Bio-Rad CFX Manager software.

RESULTS AND DISCUSSION

Sequencing and *De novo* Transcriptome Assembly

To obtain an overview of 'NAU-YH' transcriptome in taproots, and identify candidate genes involved in sucrose metabolism, a cDNA library was constructed from the RNA (an equally mixture of total RNA from three taproot developmental stages) of 'NAU-YH', and sequenced using the Illumina HiSeq™ 2000 sequencing platform. The Illumina sequencing results were shown in **Table 1**. It yielded a total of 57.0 million raw sequencing reads. After the adapter sequences, reads with unknown nucleotides larger than 5% and low quality reads were removed, 51.1 million clean pair-end reads with total of 4.6 billion nucleotides (nt) were generated for assembly. Q20 percentage, N percentage, and GC percentage were 98.29, 0.01, and 47.10%, respectively. The output was similar to previous studies on radish taproot transcriptome (Wang et al., 2012, 2013). In addition, the length of assembled sequences is an evaluation criterion for the assembly of transcriptome. In this study, the length distribution of the contigs and unigenes were shown in **Table 2**. A total of 130,953 contigs (length $\geq$ 100) were assembled with the N50 of 636 nt, an average length of 352 nt, and a total nucleotides length of 46,146,957 nt. Among them, there were 109,269 contigs (83.72%) size ranging from 100 to 500 nt, 11,699 contigs (8.93%) with size varying from 501 to 1000 nt, and 9985 contigs (7.62%) with size more than 1000 nt. Thereafter, with pair-end reads, the contigs were further generated into 70,168 unigenes with a total length of 50,277,812 nt, and with an N50 of 994 nt and a mean length of 717 nt. Meanwhile, according to a sequence similarity search with known proteins or nucleotides database, a total of 70,168 consensus sequences were assigned to 32,332 clusters and 37,846 singletons. **Table 2** also showed that the length of assembled unigenes were mostly ranged from 200 to 1000 nt accounted for 77.61%, and 15,713 unigenes (22.39%) with length > 1000 nt. These results indicated that the unigenes distribution followed the contigs distribution was greater among shorter assembled sequences.

Recently, several transcriptome studies of radish leaf and root had been reported (Wang et al., 2013; Zhang et al., 2013; Wu et al., 2015). Wu et al. (2015) reported that 68,086 unigenes with an average length of 576 bp and an N50 of 773 bp was generated from radish leaves by Trinity assembly. Wang et al. (2013) showed that 73,084 unigenes with a mean length of 763 nt and an N50 of 1095 nt were obtained from radish root transcriptome. In this study, the comparison analysis showed that the number and N50 sizes of the assembled unigenes were larger than those from the previous leaf transcriptome, while smaller than those from the root transcriptome (Wang et al., 2013; Zhang et al., 2013; Wu et al., 2015). These results implied that the quality of sequencing data was high enough to ensure the accuracy of the sequence assembly.

TABLE 1 | Output of the transcriptome sequencing for the Radish.

Feature	Statistic
Total Raw Reads	57,031,638
Total Clean Reads	51,169,168
Total Clean Nucleotides (nt)	4,605,225,120
Q20 percentage	98.29%
N percentage	0.01%
GC percentage	47.10%

TABLE 2 | Length distribution of assembled contigs and unigenes.

Nucleotides length (nt)	Contigs	Contigs percent (%)	Unigenes	Unigenes percent (%)
100–200	74216	56.67	0	0.00
201–300	19874	15.18	17228	24.55
301–400	9788	7.47	10018	14.28
401–500	5391	4.12	7006	9.98
501–600	3596	2.75	5586	7.96
601–700	2668	2.04	4540	6.47
701–800	2109	1.61	3847	5.48
801–900	1778	1.36	3306	4.71
901–1000	1548	1.18	2924	4.17
1001–1100	1297	0.99	2360	3.36
1101–1200	1098	0.84	2117	3.02
1201–1300	935	0.71	1685	2.40
1301–1400	868	0.66	1540	2.19
1401–1500	762	0.58	1319	1.88
1501–1600	659	0.50	1088	1.55
1601–1700	589	0.45	970	1.38
1701–1800	523	0.40	771	1.10
1801–1900	416	0.32	663	0.94
1901–2000	353	0.27	503	0.72
2001–2100	308	0.24	442	0.63
2101–2200	270	0.21	389	0.55
2201–2300	217	0.17	279	0.40
2301–2400	197	0.15	235	0.33
2401–2500	176	0.13	226	0.32
2501–2600	158	0.12	154	0.22
2601–2700	124	0.09	128	0.18
2701–2800	123	0.09	104	0.15
2801–2900	112	0.09	92	0.13
2901–3000	93	0.07	82	0.12
≥3000	707	0.54	566	0.81
Total number	130953		70168	
Total nucleotides length (nt)	46,146,957		50,277,812	
Mean length (nt)	352		717	
N50 (nt)	636		994	
Total consensus sequences			70168	
Distinct clusters			32332	
Distinct singletons			37846	

Functional Annotation of the Assembled Unigenes

To learn an overview information of unigene sequences from radish root transcriptome, a homology based method was

adopted in unigenes annotation. The unigene sequences were performed against public protein and nucleotide databases (Nr, Swiss-Pot, KEGG, COG and Nt) using BLAST algorithm ($E \leq 10^{-5}$) to search for sequence similarity. The results of functional annotation were shown in **Table 3**. Out of the 70,168 unigenes, 63,991 (91.20%) unigenes were matched with the public databases. The percent of annotated unigenes was similar to previously studies in radish (92.09%) by Wang et al. (2013), suggesting that the assembled unigenes have the relatively conserved functions and this project has captured the majority of the radish transcriptome. In addition, the present study 57,495 and 1384 CDS were obtained by Blast and ESTScan alignment, respectively. However, the remaining of 6177 unigene sequences, which may represent novel genes specifically expressed in radish taproot or could be attributed to other technical or biological biases such as assembly parameters, were found to be without a homologous hit in the public databases. The length distribution of CDS and predicted proteins by BLASTx and ESTScan software were shown in **Figure 1**.

For the Nr annotations, we further analyzed the *E*-value, similarity and species distribution of the top hits in the Nr database, and the results was listed in **Figure 2**. The *E*-value distribution of the top hits in the Nr database indicated that 55.97% of the mapped sequences have significant homology ($E < 1.0e^{-45}$), whereas the other 44.03% of the moderate homology sequences varied from $1.0e^{-5}$ to $1.0e^{-45}$ (**Figure 2A**). The similarity distribution displayed 56.55% of the query sequences with a similarity >80%, while 43.45% of the hits have a similarity ranging from 18 to 80% (**Figure 2B**). For the species distribution, we found that the majority of annotated sequences were similar to *Arabidopsis thaliana* (42.7%) and *A. lyrata* subsp. *Lyrata* (41.5%), followed by *Thellungiella halophile* (3.28%), *Brassica napus* (1.96%), *B. oleracea* (1.53%), *B. rapa* subsp. *Pekinensis* (1.06%), *B. rapa* (1.01%), and others (6.96%; **Figure 2C**). The BLASTx species distribution showed a bias toward *A. thaliana* and *A. lyrata subsp. Lyrata*, as well as five species with BLAST hits belonged to the Brassicaceae family, implying that the sequences of the radish transcripts obtained in the present study were assembled and annotated properly.

TABLE 3 | Summary of annotations of 'NAU-YH' unigenes in public databases.

Public database	No. of unigene hit	Percent (%)
NR	57,303	81.67
NT	62,073	88.46
Swiss-prot	36,854	52.52
KEGG	30,971	44.14
COG	17,587	25.06
GO	51,981	74.08
ALL	63,991	91.20

Functional Classification by GO and COG

Gene ontology (GO) was applied to comprehensively describe the properties of genes and their products in our transcriptome library of radish, which is an international standardized gene functional classification system. Based on the sequence similarity, 51,981 unigenes (74.08%) were categorized into 55 functional groups and summarized into three main GO categories including molecular function, cellular component and biology process (**Table 3; Figure 3**). Under the biological process category, "cellular process" (70.08%), and "metabolic process" (65.19%), were represented the most abundant of the category, suggesting that some important metabolic activities occured in root, these results were similar to previously reported study of *de novo* transcriptome analysis in radish (Wang et al., 2013) and sweet potato (Wang et al., 2010). Under the cellular component category, "cell" (91.67%) and "cell part" (91.67%) terms were prominently represented. For the category of molecular function, "binding" (51.01%) and "catalytic activity" (42.68%) were the most dominant represented terms. Moreover, only a few genes were assigned with "virion" (0.01%), "virion part" (0.01%), "protein tag" (0.01%), and "translation regulator activity" (0.01%) GO terms.

The Cluster of Orthologous Groups (COG) is a database where orthologous gene products are classified. Every protein in COG is assumed to evolve from an ancestor protein, and the whole database is built on coding proteins with complete genome as well as system evolution relationships of bacteria, algae, and eukaryotic creatures (Wang et al., 2010; Hyun et al., 2012). In this study, in order to predict and classify possible functions, the assembled unigenes were aligned to COG database. In total, 17,587 of 70,168 (25.06%) unigenes were assigned to the COG classifications (**Table 3**), which were grouped into 25 function categories (**Figure 4**). Due to some unigenes were annotated with multiple COG functions, altogether 34,972 functional annotations were generated. Among them, the five largest group included "General function prediction" (5610, 31.90%), "Transcription" (3380, 19.22%), "Replication, recombination, and repair" (2889, 16.43%), "Translation, ribosomal structure, and biogenesis" (2636, 14.99%) and "Posttranslational modification, protein turnover, chaperones" (2361, 14.96%). Conversely, five smallest groups included "Extracellular structures" (4, 0.02%), "Nuclear structure" (11, 0.06%), "RNA processing and modification" (238, 1.35%), "Nucleotide transport and metabolism" (309, 1.76%) and "Cell motility" (312, 1.77%).

Functional Classification by KEGG

Genes within the same pathway usually cooperate with each other to perform their biological function, suggesting that pathway-based analysis can help further understanding of the genes' functions (Wenping et al., 2011). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a protein database that is able to analyze gene product during metabolism process and related gene function in the cellular processes. Therefore, to identify the biological pathways being active in the taproot of radish, the assembled unigenes were mapped to KEGG protein database. Based on the sequence similarity, 30,971 unigenes could be assigned to 126 pathways (Table S2), which were grouped into five groups (**Figure 5**). These groups most represented by unigenes were metabolism (14,619 unigenes) and genetic information processing (9131 unigenes), followed by organismal systems (2155 unigenes), cellular processes (1712 unigenes) and environmental information processing

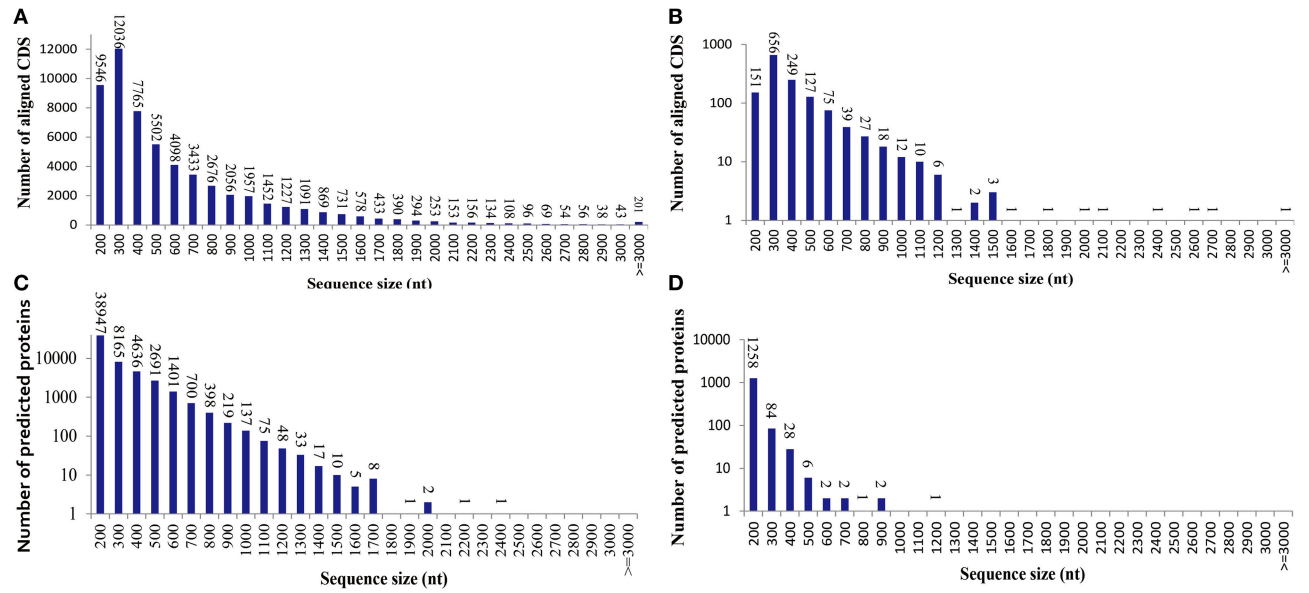


FIGURE 1 | The length distribution of the coding sequence (CDS) and predicted proteins by BLASTx and ESTScan software from the unigenes. (A) Aligned CDS by BLASTx. **(B)** Aligned CDS by ESTScan. **(C)** Predicted proteins by BLASTx. **(D)** Predicted proteins by ESTScan.

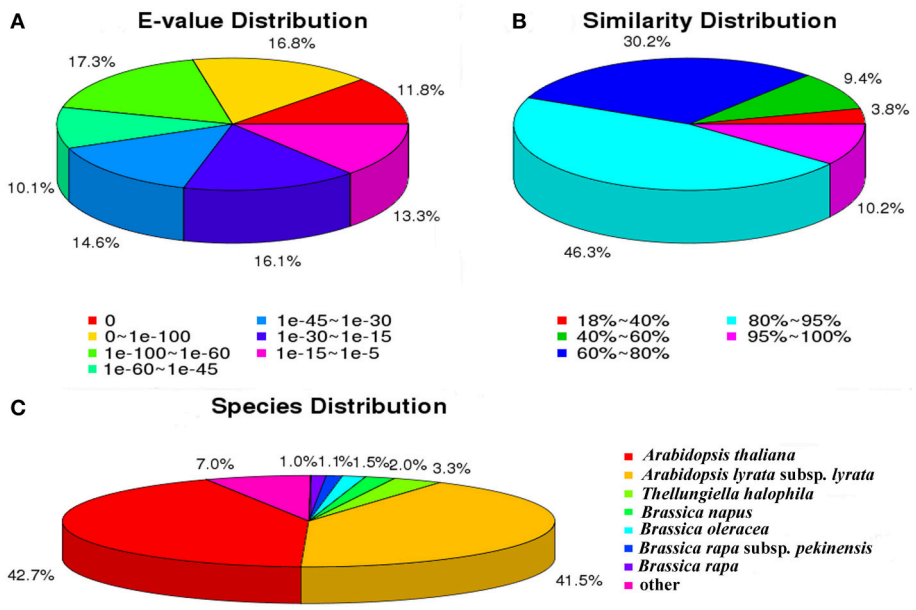


FIGURE 2 | Characteristics of sequence homology of radish root with BLAST against NCBI non-redundant (NR) database. (A) *E*-value distribution of BLAST hits for matched unigene sequences, using an *E*-value cutoff of 10^{-5}. **(B)** Similarity distribution of top BLAST hits for each unigene. **(C)** Species distribution of the top BLAST hits.

(770 unigenes). In metabolism pathway (**Figure 5**), 14,619 unigenes were divided into 10 sub-categories, of which most representation by unigenes were carbohydrate metabolism (3725 unigenes), lipid metabolism (2613 unigenes), amino acid metabolism (1919 unigenes), biosynthesis of other secondary metabolites (1295 unigenes), and nucleotide metabolism (1239 unigenes). Taken together, the putative KEGG pathways identified in the present study elucidated specific responses and functions involved in the molecular processes of radish taproot development, and provided a resource for further investigating specific pathways in radish including the sucrose metabolism pathway.

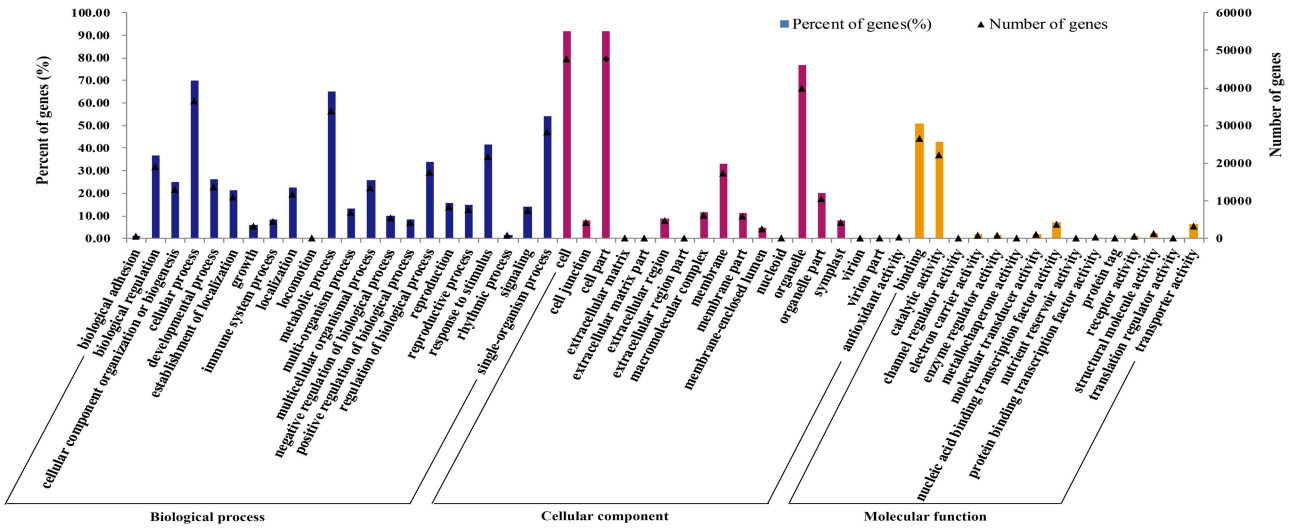


FIGURE 3 | Frequencies and mean expression levels of transcripts matching GO terms. The percentage of transcripts matching GO terms is showing for each category as different color bars and the normalized mean expression levels of transcripts matching each of these GO terms are shown as black triangles.

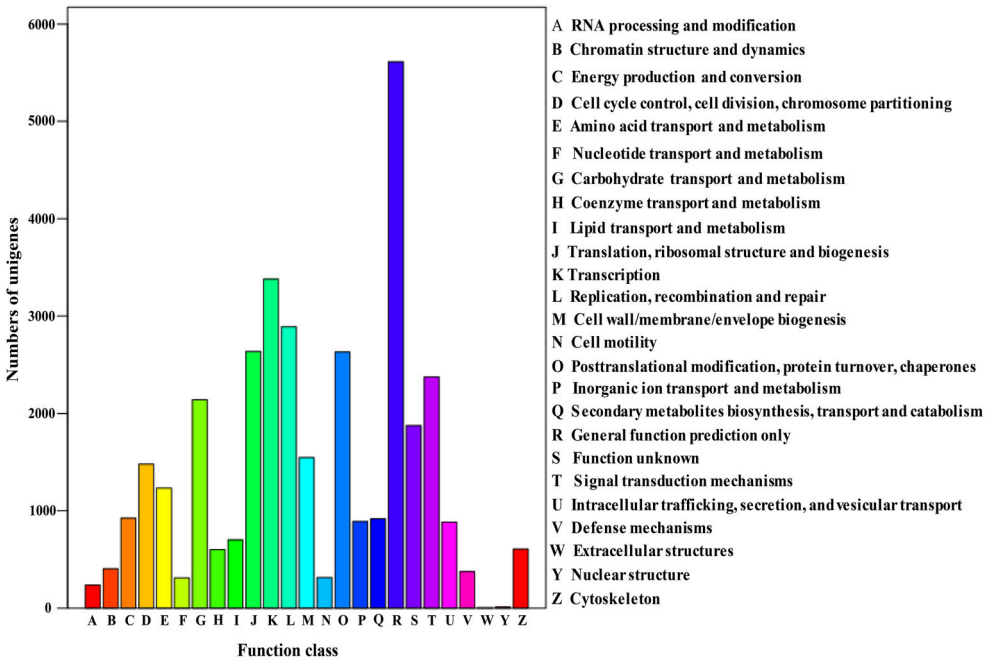


FIGURE 4 | Histogram presentation of clusters of orthologous groups (COG) classification.

Analysis of Sucrose Metabolism Pathway Genes Using Radish Unigenes

Sucrose is the major product of photosynthesis, and it is the main substrate for sink strength, and used to sustain cell metabolism and growth (Tognetti et al., 2013; Ruan, 2014). In radish, the storage root is a major sink, which begins to thicken early in development (Usuda et al., 1999). To date, the main pathway of sucrose metabolism has been well-known in higher plant (Ruan, 2012, 2014; Zhang et al., 2015). In our annotated radish taproot transcriptome dataset, a total of 103 transcripts encoding eight well-known enzymes involved in the main sucrose metabolism pathway were identified by KEEG protein database (**Figure 6**). Transcript IDs from the sucrose metabolism pathway were listed in Table S3A. The sucrose biosynthesis in cytosol has been proposed by two key enzymes: SPS (EC 2.4.1.14, 16 transcripts) and Sucrose-phosphate phosphatase

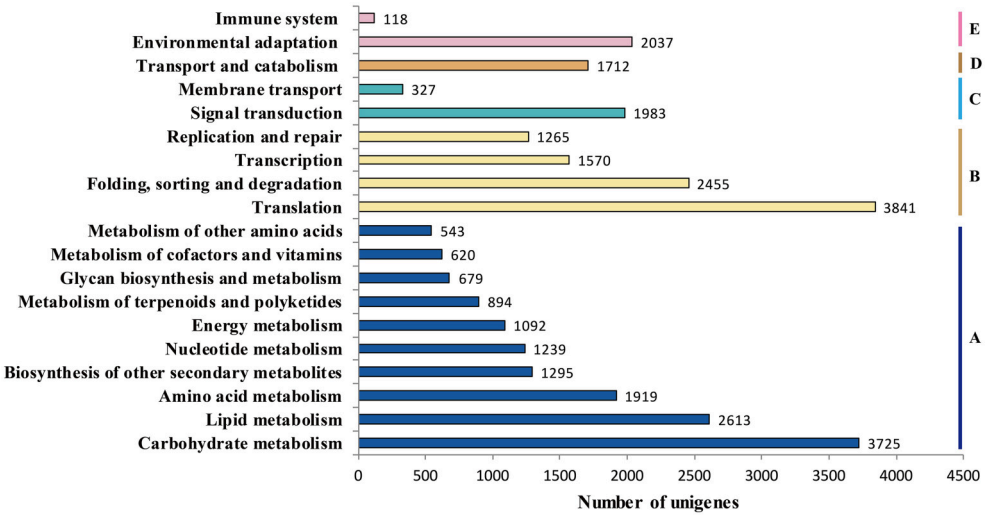


FIGURE 5 | Pathway assignment based on KEGG. (A) metabolism; **(B)** genetic information processing; **(C)** environmental information processing; **(D)** cellular processes; **(E)** environmental adaptation.

(SPP; EC 3.1.3.24, no annotated transcripts available by KEEG protein database) (Ruan, 2012; Zhang et al., 2015), that is, Glucose (Glc) use hexokinase (EC 2.7.1.1, 11 transcripts) as substrates to generate Glc-6-phosphate (Glc-6-P), which can be converted to fructose-6-phosphate (Fru-6-P) by Glc-6-P isomerase (EC.5.3.1.9, five transcripts). Following this reaction, SPS uses Fru-6-P and UDP-Glc as substrates to produce Sucrose-phosphate (Sucrose-6-P), which is then converted to sucrose by sucrose-phosphatase or sucrose-6-phosphate phosphohydrolase (SPP; EC 3.1.3.24, six transcripts by Nr annotation and BLASTx manual) (Table S3A; **Table 4**). Furthermore, evidence shows that the import and accumulation of sucrose in storage roots might involve its inversion into hexose sugars for use in diverse ways by invertase and sucrose synthase (Ruan, 2014). In this study, many transcripts encoding critical functional enzyme involved in two possible sucrose degradation pathways were also discovered in our transcriptome. One is the conversion of sucrose to glucose and fructose by invertase or beta-fructofuranosidase (EC 3.2.1.26, 17 transcripts). Another is the conversion of sucrose to UDP-Glucose and fructose by SuSy (EC 2.4.1.13, 29 transcripts; Table S3A; **Table 4**; **Figure 6**). In addition, as shown in **Table 4**, SuSy, INV, and SPS were encoded by the high numbers of transcripts, implying that these enzymes are the major source of sucrose metabolism activity in radish (Ren and Zhang, 2013). Moreover, significant numbers of transcripts for fructokinase metabolism may represent its property of the taproot having a sweet taste (Zhang et al., 2015).

To investigate which transcripts were unique involved in the main sucrose metabolism pathway in annotated 'NAU-YH' taproot transcriptome dataset, the transcripts encoding eight well-known enzymes involved in the main sucrose metabolism pathway by KEEG protein database were annotated from 'NAU-RG' taproot transcriptome dataset available in our lab [SRX316199 and http://www.ncbi.nlm.nih.gov/sra/] (Wang et al., 2013). A total of 127 transcripts were annotated in 'NAU-RG' taproot transcriptome dataset (Table S3B). Among of these, SuSy, INV, and SPS were also encoded by the higher numbers of transcripts. In addition, the 'NAU-YH' transcripts encoding eight well-known enzymes in the main sucrose metabolism pathway by KEEG protein database (Table S3A) were compared to the transcripts of 'NAU-RG' (Table S3B) by using local BLASTN with an *E*-value cutoff of $1e^{-20}$ (Table S3C). As a result, all transcripts (Table S3A) in 'NAU-YH' showed significant identity to the transcripts of 'NAU-RG' (Table S3C). These results indicated that the transcripts encoding enzymes in the main sucrose metabolism pathway were similar in these two radish genotypes.

Validation and Expression Analysis of Genes Involved in Sucrose Metabolism

To assess the quality of the assembly and annotation data from radish taproot transcriptome sequencing, full-length cDNA sequences of two key genes from sucrose metabolism process were isolated by T-A cloning with the Sanger method and compared with the assembled sequences. The length of *RsSPS1* and *RsSuSy*1 genes were 3265 and 2163 bp, respectively (**Table 5**). Overall, the assembled unigenes covered 92.75% (*RsSPS1*) and 98.28% (*RsSuSy*1) of the corresponding full-length genes. Additionally, *RsSPS1*, and *RsSuSy*1 genes were predicted to contain the complete ORF, and the ORF pairwise identity of *RsSPS1* and *RsSuSy*1were 96.99 and 98.75%, respectively (**Table 5**). These results validated that the NGS-based RNA-seq procedures was reliable (**Table 5**).

To experimentally confirm that the unigenes obtained from sequencing and computational analysis were indeed expressed, six unigenes related to sucrose metabolism pathway were chosen for RT-qPCR analysis (**Figure 7**). The RT-qPCR analysis showed that all the genes exhibited different expression and regulation during radish taproot thickening. The expression profiles of

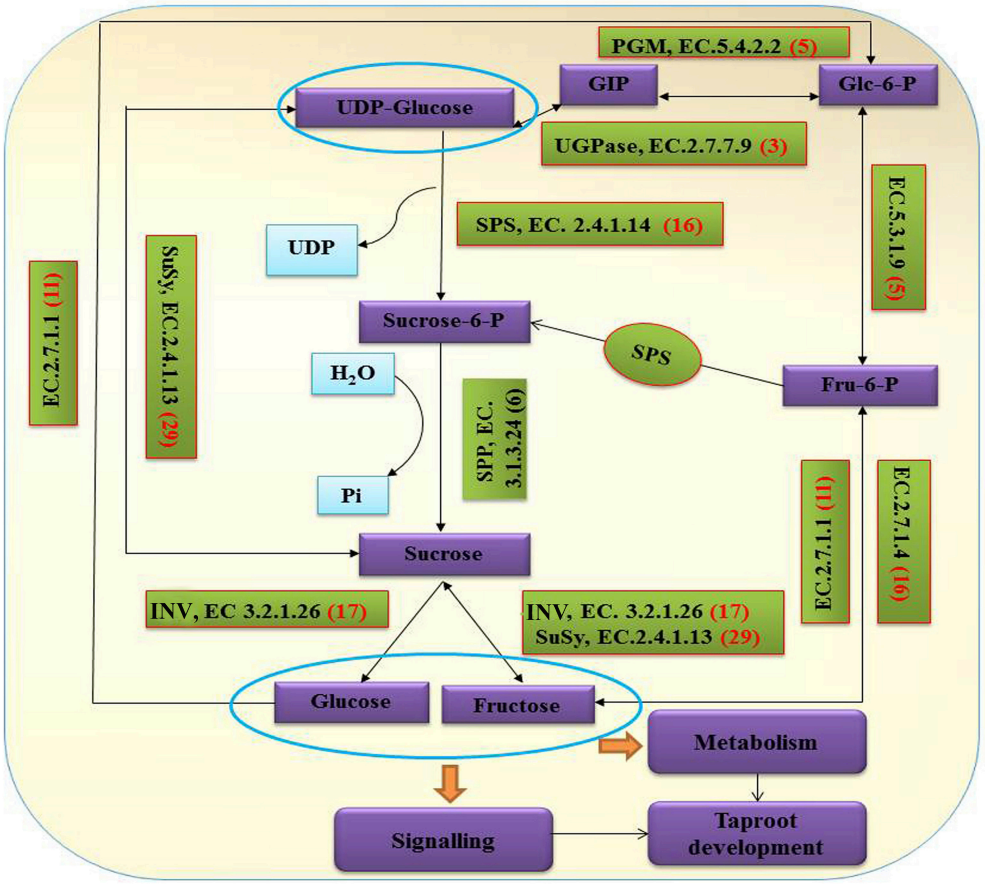


FIGURE 6 | Assembled radish unigenes that may be involved in the sucrose metabolism pathway. The numbers in brackets following each gene name indicate the number of transcriptome unigenes annotated to that gene.

TABLE 4 | Classification analysis of unigenes coding for 8 protein families in sucrose metabolism in radish taproot transcriptomes.

Family	KEGG annotation [international enzyme name]	Unigene no.[b]	NAU-YH_raw fragments (no.)[c]	NAU-YH_RPKM[d]
1[a]	sucrose-phosphate synthase [EC:2.4.1.14]	16	6254	195.175
2[a]	sucrose synthase [EC:2.4.1.13]	29	24824	1072.2468
3[a]	invertase (beta-fructofuranosidase) [EC:3.2.1.26]	17	1609	134.0041
4[a]	hexokinase [EC:2.7.1.1]	11	2290	102.1245
5[a]	UTP–glucose-1-phosphate uridylyltransferase [EC:2.7.7.9]	3	8622	278.7609
6[a]	phosphoglucomutase [EC:5.4.2.2]	5	4006	141.776
7[a]	glucose-6-phosphate isomerase [EC:5.3.1.9]	5	2838	112.6653
8[a]	fructokinase [EC:2.7.1.4]	17	13035	531.1143
9	sucrose-phosphatase or sucrose-6-phosphate phosphohydrolase (EC 3.1.3.24)	6	1381	67.7794

[a] *Gene families from 1 to 8 were all identified from "starch and sucrose metabolism" category according to the KEGG protein database;*

[b] *The numbers under "Unigene no." column represent the total number of unigenes in each enzyme family identified in the radish taproot transcriptome;*

[c] *The numbers under the "NAU-YH_raw fragments (no.)" column represent the raw fragments in each enzyme family identified in the radish taproot transcriptome;*

[d] *The numbers under "RPKM" columns represent the total values of unigene RPKM in each enzyme families identified in the radish taproot transcriptome.*

ATBFRUCT1 and *cwINV6* were similar in radish different tissues and different development stages. And they were highly expressed in the root and leaf, especially in root organ of pre-cortex splitting stage (10 DAS). *SUS1* had higher expression profiles in root organ during the different taproot thickening stages. The highest expression level of *SUS1* was observed in root organ of expanding stage (40 DAS). *SUS3* exhibited higher expression in root and leaf from cortex splitting stage (20 DAS) to mature stage (50 DAS), the highest expression level in root organ at cortex splitting stage, whereas higher expression was

observed in leaf at mature stage. The results were consistent with previous studies (Usuda et al., 1999; Rouhier and Usuda, 2001), suggesting they may be involved in radish taproot formation. *SPS1* was highly expressed in stem at mature stage. In contrast, the expression levels of *SPS2* in leaves were higher in roots and stems during the taproot formation, and the highest expression level of *SPS2* was observed at cortex splitting stage.

TABLE 5 | Sequence analyses of the two putative radish genes involved in sucrose metabolism process.

Gene	Full-length cDNA (bp)	Unigene No.	Coverage (%)	ORF similarity (%)
RsSPS1	3265	9	92.75	96.99
RsSuSy1	2163	13	98.28	98.75

CONCLUSION

In summary, a cDNA library was sequenced using NGS-based Illumina sequencing platform. From ∼51 million clean reads, a total of 70,168 unigene with a total length of 50.28 Mb, an average length of 717 bp and a N50 of 994 bp were obtained. In total, 63,991 (about 91.20% of the assembled unigenes) unigenes were successfully annotated to five public databases including NR, GO, COG, KEGG, and Nt. GO term analysis revealed that the

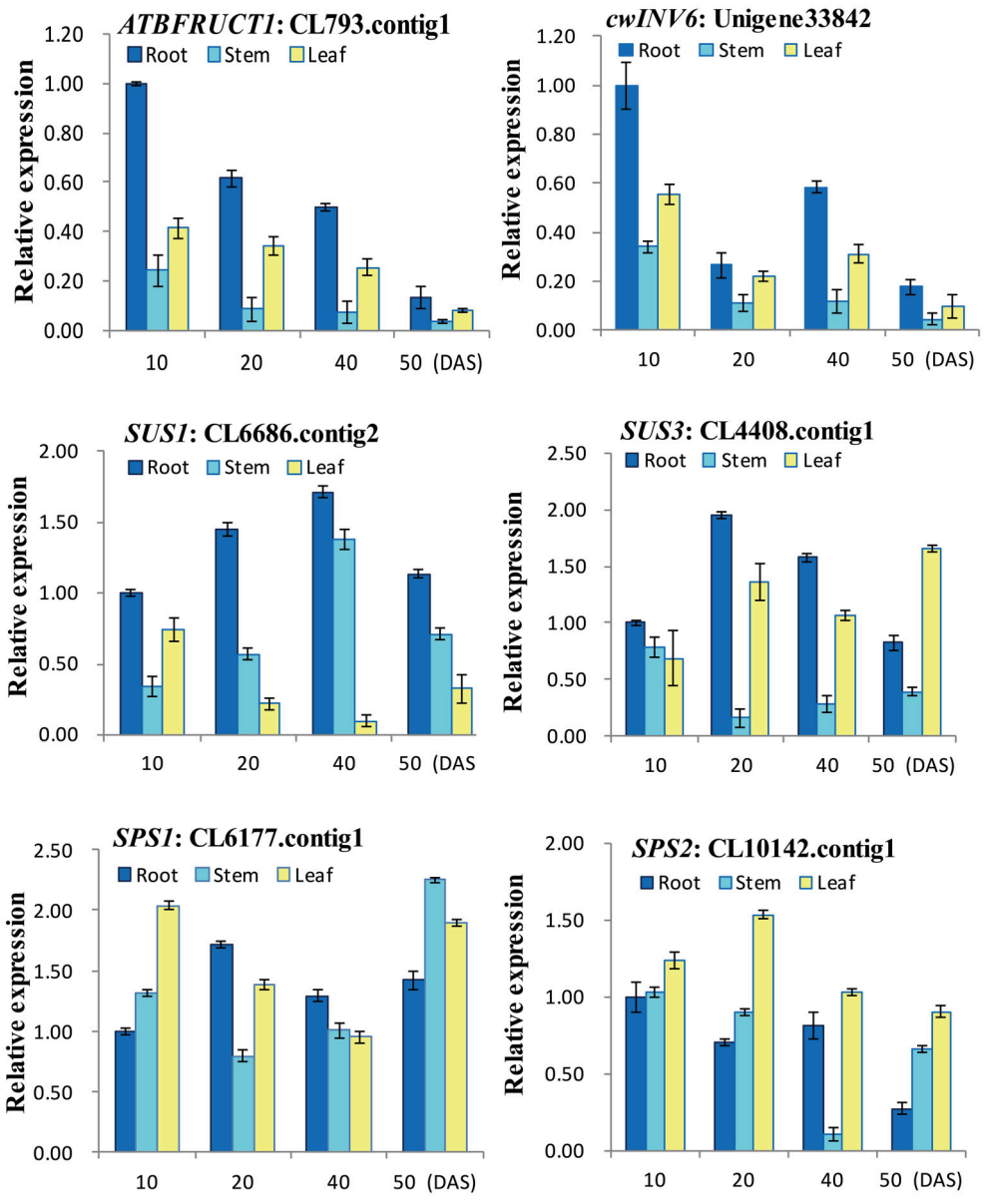


FIGURE 7 | RT-qPCR analysis of six genes involved in sucrose metabolism with different tissues and stages in radish.

majority of these unigenes were predominately involved in basic physiological and metabolic processes, catalytic, binding and cellular process. Furthermore, a total of 103 unigenes encoding eight enzymes in the sucrose metabolism related pathways were identified. These results provided an solid foundation for identifying taproot thickening-related critical genes and would facilitate further dissecting molecular mechanisms underlying taproot formation in radish.

ACCESSION CODE

The RNA SEQ raw data of this study have been deposited in NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra) with accession number: SRX707630.

AUTHOR CONTRIBUTIONS

YR and LW designed the experiments. YR, XL and WY performed the radish cultivation and sample collection. YR, WR, ZW and XY performed the experiments. YR wrote the manuscript draft. LW, ZX, KB and XL edited and revised the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was in part supported by grants from the NSFC (31372064, 31501759, 30571193), Key Technology R & D Program of Jiangsu Province (BE2013429), JASTIF and the PAPD.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00585

REFERENCES

Angeloni, F., Wagemaker, C. A., Jetten, M. S., Op den Camp, H. J., Janssen megens, E. M., Francoijs, K. J., et al. (2011). *De novo* transcriptome characterization and development of genomic tools for *Scabiosa columbaria* L. using next-generation sequencing techniques. *Mol. Ecol. Resour.* 11, 662–674. doi: 10.1111/j.1755-0998.2011.02990.x

Babb, V. M., and Haigler, C. H. (2001). Sucrose phosphate synthase activity rises in correlation with high-rate cellulose synthesis in three heterotrophic systems. *Plant Physiol.* 127, 1234–1242. doi: 10.1104/pp.010424

Chaturvedi, P. (2008). Inhibitory response of *Raphanus sativus* on lipid peroxidation in albino rats. *Evid. Based Complement. Alternat. Med.* 5, 55–59. doi: 10.1093/ecam/nel077

Cheng, W. H., and Chourey, P. S. (1999). Genetic evidence that invertase-mediated release of hexoses is critical for appropriate carbon partitioning and normal seed development in maize. *Theor. Appl. Genet.* 98, 485–495. doi: 10.1007/s001220051096

Conesa, A., Götz, S., García Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Curtis, I. S. (2003). The noble radish: past, present and future. *Trends Plant Sci.* 8, 305–307. doi: 10.1016/S1360-1385(03)00127-4

Farrar, J., Pollock, C., and Gallagher, J. (2000). Sucrose and the integration of metabolism in vascular plants. *Plant Sci.* 154, 1–11. doi: 10.1016/S0168-9452(99)00260-5

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Gutiérrez, R. M., and Perez, R. L. (2004). *Raphanus sativus* (Radish): their chemistry and biology. *Scientific World Journal* 4, 811–837. doi: 10.1100/tsw.2004.131

Hyun, T. K., Rim, Y., Jang, H. J., Kim, C. H., Park, J., Kumar, R., et al. (2012). *De novo* transcriptome sequencing of Momordica cochinchinensis to identify genes involved in the carotenoid biosynthesis. *Plant Mol. Biol.* 79, 413–427. doi: 10.1007/s11103-012-9919-9

Iseli, C., Jongeneel, C. V., and Bucher, P. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 138–148.

Johnston, J. S., Pepper, A. E., Hall, A. E., Chen, Z. J., Hodnett, G., Drabek, J., et al. (2005). Evolution of genome size in *Brassicaceae*. *Ann. Bot.* 95, 229–235. doi: 10.1093/aob/mci016

Kitashiba, H., Li, F., Hirakawa, H., Kawanabe, T., Zou, Z., Hasegawa, Y., et al. (2014). Draft sequences of the radish (*Raphanus sativus* L.) genome. *DNA Res.* 21, 481–490. doi: 10.1093/dnares/dsu014

Li, X. Q., and Zhang, D. (2003). Gene expression activity and pathway selection for sucrose metabolism in developing storage root of sweet potato. *Plant Cell Physiol.* 44, 630–636. doi: 10.1093/pcp/pcg080

Ren, X., and Zhang, J. (2013). Research progresses on the key enzymes involved in sucrose metabolism in maize. *Carbohydr. Res.* 368, 29–34. doi: 10.1016/j.carres.2012.10.016

Rouhier, H., and Usuda, H. (2001). Spatial and temporal distribution of sucrose synthase in the radish hypocotyl in relation to thickening growth. *Plant Cell Physiol.* 42, 583–593. doi: 10.1093/pcp/pce071

Ruan, Y. L. (2012). Signaling role of sucrose metabolism in development. *Mol. Plant* 5, 763–765. doi: 10.1093/mp/sss046

Ruan, Y. L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu. Rev. Plant Biol.* 65, 33–67. doi: 10.1146/annurev-arplant-050213-040251

Stokes, M. E., Chattopadhyay, A., Wilkins, O., Nambara, E., and Campbell, M. M. (2013). Interplay between sucrose and folate modulates auxin signaling in *Arabidopsis*. *Plant Physiol.* 162, 1552–1565. doi: 10.1104/pp.113.215095

Tognetti, J. A., Pontis, H. G., and Martínez Noël, G. M. (2013). Sucrose signaling in plants: a world yet to be explored. *Plant Signal. Behav.* 8:e23316. doi: 10.4161/psb.23316

Usuda, H., Demura, T., Shimogawara, K., and Fukuda, H. (1999). Development of sink capacity of the "storage root" in a radish cultivar with a high ratio of "storage root" to shoot. *Plant Cell Physiol.* 40, 369–377. doi: 10.1093/oxfordjournals.pcp.a029552

Wang, F., Sanz, A., Brenner, M. L., and Smith, A. (1993). Sucrose synthase, starch accumulation, and tomato fruit sink strength. *Plant Physiol.* 101, 321–327.

Wang, L., and He, Q. (2005). *Chinese Radish*. Beijing: Scientific and Technical Documents Publishing House.

Wang, S., Wang, X., He, Q., Liu, X., Xu, W., Li, L., et al. (2012). Transcriptome analysis of the roots at early and late seedling stages using Illumina paired-end sequencing and development of EST-SSR markers in radish. *Plant Cell Rep.* 31, 1437–1447. doi: 10.1007/s00299-012-1259-3

Wang, W., Gong, Y., Liu, L., Wang, Y., Jing, Z., Huang, D., et al. (2007). Changes of sugar content and sucrose metabolizing enzyme activities during fleshy taproot development in radish (*Raphanus sativus* L.). *Acta Hortic. Sin.* 34, 1313–1316. doi: 10.3321/j.issn:0513-353x.2007.05.043

Wang, Y., Pan, Y., Liu, Z., Zhu, X., Zhai, L., Xu, L., et al. (2013). *De novo* transcriptome sequencing of radish (*Raphanus sativus* L.) and analysis of major genes involved in glucosinolate metabolism. *BMC Genomics* 14:836. doi: 10.1186/1471-2164-14-836

Wang, Z., Fang, B., Chen, J., Zhang, X., Luo, Z., Huang, L., et al. (2010). *De novo* assembly and characterization of root transcriptome using Illumina paired-end sequencing and development of cSSR markers in sweetpotato (*Ipomoea batatas*). *BMC Genomics* 11:726. doi: 10.1186/1471-2164-11-726

Ward, J. A., Ponnala, L., and Weber, C. A. (2012). Strategies for transcriptome analysis in nonmodel plants. *Am. J. Bot.* 99, 267–276. doi: 10.3732/ajb.1100334

Weber, H., Borisjuk, L., and Wobus, U. (1997). Sugar import and metabolism during seed development. *Trends Plant Sci.* 2, 169–174. doi: 10.1016/S1360-1385(97)85222-3

Wenping, H., Yuan, Z., Jie, S., Lijun, Z., and Zhezhi, W. (2011). *De novo* transcriptome sequencing in *Salvia miltiorrhiza* to identify genes involved in the biosynthesis of active ingredients. *Genomics* 98, 272–279. doi: 10.1016/j.ygeno.2011.03.012

Wu, G., Zhang, L., Yin, Y., Wu, J., Yu, L., Zhou, Y., et al. (2015). Sequencing, *de novo* assembly and comparative analysis of *Raphanus sativus* transcriptome. *Front. Plant Sci.* 6:198. doi: 10.3389/fpls.2015.00198

Xiong, Y., McCormack, M., Li, L., Hall, Q., Xiang, C., and Sheen, J. (2013). Glucose-TOR signalling reprograms the transcriptome and activates meristems. *Nature* 496, 181–186. doi: 10.1038/nature12030

Xu, Y., Zhu, X., Gong, Y., Xu, L., Wang, Y., Liu, L., et al. (2012). Evaluation of reference genes for gene expression studies in radish (*Raphanus sativus* L.) using quantitative real-time PCR. *Biochem. Biophys. Res. Commun.* 3, 398–403. doi: 10.1016/j.bbrc.2012.06.119

Yang, L., Xu, M., Koo, Y., He, J., and Poethig, R. S. (2013). Sugar promotes vegetative phase change in *Arabidopsis thaliana* by repressing the expression of MIR156A and MIR156C. *Elife* 2:e00260. doi: 10.7554/eLife.00260

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, W293–W297. doi: 10.1093/nar/gkl031

Zhang, L., Jia, H., Yin, Y., Wu, G., Xia, H., Wang, X., et al. (2013). Transcriptome analysis of leaf tissue of *Raphanus sativus* by RNA sequencing. *PLoS ONE* 8:e80350. doi: 10.1371/journal.pone.0080350

Zhang, L., Lin, Q., Feng, Y., Fan, X., Zou, F., Yuan, D. Y., et al. (2015). Transcriptomic identification and expression of starch and sucrose metabolism genes in the seeds of Chinese Chestnut (*Castanea mollissima*). *J. Agric. Food Chem.* 63, 929–942. doi: 10.1021/jf505247d

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Yu, Xu, Zhang, Wang, Luo, Wang, Zhu, Xie, Karanja and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 July 2016
doi: 10.3389/fpls.2016.01012

Modification of Leaf Glucosinolate Contents in *Brassica oleracea* by Divergent Selection and Effect on Expression of Genes Controlling Glucosinolate Pathway

*Tamara Sotelo, Pablo Velasco, Pilar Soengas, Víctor M. Rodríguez and María E. Cartea **

Group of Genetics, Breeding and Biochemistry of Brassicas, Misión Biológica de Galicia-Consejo Superior de Investigaciones Científicas, Pontevedra, Spain

OPEN ACCESS

Edited by:
Juan Francisco Jimenez Bremont, Instituto Potosino de Investigación Científica y Tecnológica, Mexico

Reviewed by:
Zhongyun Piao, Shenyang Agricultural University, China
Hao Peng, Washington State University, USA

***Correspondence:**
María E. Cartea
ecartea@mbg.csic.es

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *19 February 2016*
Accepted: *27 June 2016*
Published: *15 July 2016*

Citation:
Sotelo T, Velasco P, Soengas P, Rodríguez VM and Cartea ME (2016) Modification of Leaf Glucosinolate Contents in Brassica oleracea by Divergent Selection and Effect on Expression of Genes Controlling Glucosinolate Pathway. Front. Plant Sci. 7:1012. doi: 10.3389/fpls.2016.01012

Modification of the content of secondary metabolites opens the possibility of obtaining vegetables enriched in these compounds related to plant defense and human health. We report the first results of a divergent selection for glucosinolate (GSL) content of the three major GSL in leaves: sinigrin (SIN), glucoiberin (GIB), and glucobrassicin (GBS) in order to develop six kale genotypes (*Brassica oleracea* var. *acephala*) with high (HSIN, HIGIB, HGBS) and low (LSIN, LGIB, LGBS) content. The aims were to determine if the three divergent selections were successful in leaves, how each divergent selection affected the content of the same GSLs in flower buds and seeds and to determine which genes would be involved in the modification of the content of the three GSL studied. The content of SIN and GIB after three cycles of divergent selection increased 52.5% and 77.68%, and decreased 51.9% and 45.33%, respectively. The divergent selection for GBS content was only successful and significant for decreasing the concentration, with a reduction of 39.04%. Mass selection is an efficient way of modifying the concentration of individual GSLs. Divergent selections realized in leaves had a side effect in the GSL contents of flower buds and seeds due to the novo synthesis in these organs and/or translocation from leaves. The results obtained suggest that modification in the SIN and GIB concentration by selection is related to the GSL-ALK locus. We suggest that this locus could be related with the indirect response found in the GBS concentration. Meantime, variations in the CYP81F2 gene expression could be the responsible of the variations in GBS content. The genotypes obtained in this study can be used as valuable materials for undertaking basic studies about the biological effects of the major GSLs present in kales.

Keywords: divergent mass selection, glucosinolates, *Brassica oleracea*, GSL-ALK, RT-qPCR, modifying gene expression

INTRODUCTION

Glucosinolates (GSLs) are a major class of secondary metabolites found in the family *Brassicaceae*. Due to their enhanced plant protection to biotic and abiotic stresses (Fahey et al., 2001; Santolamazza–Carbone et al., 2014) and their preventive effects on several human cancers (Fahey and Stephenson, 1999; Forte et al., 2008), they have been extensively investigated. The hydrolytic breakdown products of GSLs, especially isothiocyanates (ITCs), have beneficial effects on human health, such as cytotoxic and apoptotic effects in damaged cells, preventing cancer in humans and reducing risk for degenerative diseases (Cartea and Velasco, 2008; Forte et al., 2008; Van Horn et al., 2008; Virgili and Marino, 2008). In contrast, in rapeseed meal, the dominant GSL, progoitrin (2-hidroxy-3-butenyl GSL, PRO) is changed into an oxazolidine-2-thione, which causes goiter and has other detrimental effects on animal health (Liu et al., 2012).

GSLs are sulfur-rich plant secondary metabolites with a basic skeleton consisting of a β-thioglucose residue, an N-hydroxy monosulfate moiety, and a variable side chain (Halkier and Du, 1997; Kliebenstein et al., 2001b). Generally, GSLs are divided into three different classes according to the amino acid precursor in biosynthesis and are called aromatic GSLs (derived from phenylalanine or tyrosine), aliphatic GSLs (derived from methionine, alanine, valine, leucine, and isoleucine) and indolic GSLs (synthesized from tryptophan) (Zukalova and Vasak, 2002; Bekaert et al., 2012).

GSL biosynthesis is a tripartite pathway involving three independent steps: (i) side chain elongation, which is carried out by methylthioalkylmalate synthase enzymes (MAM). (ii) Development of the core structure, which includes several steps: aldoxime formation catalyzed by the CYP79 family of cytochromes P450; aldoxime oxidation by the CYP83 family; thiohydroximic acid formation by conjugation to an S donor and after C-S bond cleavage; desulfoGSL formation by S-glucosyltransferase (S-GT); and GSL formation by sulfotransferase. (iii) Secondary modification of the amino acid side chain, which includes oxidation, hydroxylation, methoxylation, desaturation, sulfation and glycosylation (Sorensen, 1988; Mikkelsen et al., 2002). Side chain elongation and development of the core structure are common to the three types of GSLs biosynthesis (**Figure S1**).

It is known that three loci mainly determine the profile and content of aliphatic GSLs in *B. oleracea*. The presence of 3C-GSL is controlled by a dominant allele of GSL-PRO whereas the presence of 4C-GSL and 5C-GSL is controlled by a dominant allele of GSL-ELONG. Another major gene involved in the synthesis of aliphatic GSL is GSL-ALK, which controls the conversion of methylsulphinyl GSL into alkenyl GSL (Li et al., 2001), a step related with the production of sinigrin (2-propenyl, SIN) and gluconapin (3-butenyl, GNA). The indolic GSLs pathway has been less studied than the aliphatic GSLs one. There are key loci that synthesized the core structure of indolic GSLs biosynthesis such as CYP79B2, CYP79B3, or CYP83B1 (Mikkelsen et al., 2000; Bak et al., 2001; Naur et al., 2003).

The increase of beneficial GSLs and the reduction of detrimental GSLs are a target in brassica improvement in order to obtain crops with high value and improved food quality. On the other hand, the obtaining of plant material with the same genetic background but with different concentrations of specific GSLs will allow us to study their biological effects. The first modification of GSLs content by classical breeding took place in the 70s, when low erucic acid and low GSLs content varieties of *B. napus* were obtained by introgression from other *B. napus* cultivars (Stefansson and Kondra, 1975; Röbbelen and Thies, 1980). In the 90s, UK groups held a screening of diverse wild *Brassica* species and found that *Brassica villosa* contained a high concentration of glucoraphanin (4-methylsulphinylbutyl, GRA). This wild species was crossed with a commercial broccoli leading to the production of a new cultivar of broccoli enriched in GRA (Mithen et al., 2003; Sarikamis et al., 2006). More recently, molecular biology techniques were applied to modify the content of a particular GSL. Liu et al. (2012) obtained *B. napus* seeds enriched in GRA though the GSL-ALK silencing using RNAi.

The accumulation and profile of GSLs are highly dependent on the genotype, although they are also affected by environmental and developmental factors (Kliebenstein et al., 2001a; Brown et al., 2003). The concentration of GSLs shows a high variability among species, different varieties of the same species or even among plants of the same variety (Kushad et al., 1999). Divergent mass selection has been widely used in plant breeding as it can generate groups of individuals that share the same genetic background but with extreme values for a particular trait. Stowe and Marquis (2011) used this type of selection to effectively modify the total GSLs content of leaves of a rapid cycling variety of *B. rapa*. This kind of selection could also be used to modify the content of a particular GSL. The content and profile of these secondary metabolites vary with plant organs (Brown et al., 2003; Velasco et al., 2007). Selection carried out in one organ could produce side effects on the content of GSLs in other organs of the plant. Modification or selection by one gene of the GSLs biosynthetic pathway can also produce alterations or modifications in the concentration of other GSLs within the same biosynthetic pathway. On the other hand, it would be interesting to know the action of which genes are being modified in the process of divergent selection. This information will allow us to select directly for those genes that are involved in this change.

In kales (*Brassica oleracea* var. *acephala*), two aliphatic GSLs, SIN and glucoiberin (3-methylsulphinylpropyl, GIB), and one indolic GSL, glucobrassicin (3-indolylmethyl, GBS), are the predominant in the leaf profile (Velasco et al., 2007; Cartea and Velasco, 2008). We report herein the results of three cycles of divergent mass selection for GIB, SIN and GBS content in leaves. This on-going selection program provides unique germplasm to study the direct and indirect effects of selection on individual GSLs concentration. Our objectives were: (1) studying the effect to the divergent selections for the content of two aliphatic GSLs (GIB and SIN) and one indolic GSL (GBS) in leaves, (2) determining the side effect of divergent selections in seeds and flower buds, (3) establishing whether the content of other GSLs may be altered with the selections carried on in leaves and (4)

determining which genes would be involved in the modification of the leaf content of GIB, SIN, and GBS in divergent selections.

MATERIALS AND METHODS

Divergent Selection Program

Divergent selections were started in 2006 by using seeds of the kale population MBG-BRS0062, kept at the brassica germplasm bank at Misión Biológica de Galicia (MBG-CSIC) (Galicia, NW Spain). The population presents variability for GSL concentration and this is a desirable characteristic to realize a mass divergent selection for high and low content. Divergent selections were designed to obtain plant varieties with high (HSIN) and low (LSIN) SIN content, high (HGIB) and low (LGIB) GIB content, and high (HGBS) and low (LGBS) GBS content. In 2006, approximately 750 plants from cycle 0 (C0) were transplanted in the field into six cages (125 plants each for each one of the selections, i.e., HSIN, LSIN, HGIB, LGIB, HGBS, and LGBS. The leaf GSL content of all the plants was assessed 120 days after sowing by UHPLC. After analysis, the 25 plants (≈20% selection intensity) with the highest content were selected in HSIN, HGIB and HGBS, and the 25 plants with the lowest content were selected in LSIN, LGIB and LGBS. Non-selected plants were removed from the cages before flowering. Cross-pollination among the selected plants in each cage was made by bumblebees (*Bombus terrestris*). Afterwards, seed of the selected plants were mixed and in this way the cycle 1 (C1) of each one of the selections was obtained. From 2008 to 2009, this process was repeated to obtain the cycles C2 and C3, respectively for all the high and low selections. After finishing the process of selection, and in order to recombine each one of the genotypes and to obtain seed for each selection cycles in the same environmental conditions, all the selection cycles plus the original population (C0, C1, C2, and C3 for each SIN, GIB, and GBS) were multiplied in 2010 in isolated experimental plots at MBG-CSIC.

Evaluation Trials

Two different assays were carried out. A field trial was conducted to test the effectiveness of divergent selection. Recombined plants from 18 cycles of divergent selection (C1, C2, and C3 for HSIN, LSIN, HGIB, LGIB, HGBS, and LGBS) plus the original cycle (C0) were studied in the same year in order to avoid variations on GSLs content due to environmental conditions. The study was conducted during 2012 at MBG-CSIC (Galicia, NW Spain). Plants were grown in multi-pot trays under controlled conditions in an acclimatized greenhouse from July to August in 2012. On 29th August plants were transplanted into the field (Salcedo, NW Spain, 42° 24′N, 8° 38′W) at 5–6 true leaf stages. Experimental design was a randomized complete block with three replicates. Each plot had two rows spaced 0.8 m and each row consisted of 15 plants spaced 0.6 m.

The evaluation of C0 with the same precision than the other cycles requires a considerably larger number of experimental plots, as this population contained 100% of the initial variability for GSL concentration. For this reason, three plots of the C0 were planted per block, while for the other genotypes one plot per block was planted. This variability was of less magnitude in the rest of cycles, because their starting variability had been reduced by the first cycle of selection. Cultivation operations, fertilization, and weed control were carried out according to local practices and crop requirements. Leaf samples were harvested on ≈90 days old plants. The third leaf of a total of 20 healthy and competitive plants from each plot was chosen as plant material for GSLs analysis. Leaf samples were divided in two different bulks. Flower bud samples were collected from the same experimental plot sequentially, depending on the flowering time of each variety. In this case, 15 flower buds were collected and divided in three bulks from each plot. Tissue samples from leaves and flower buds were stored at −80°C, freeze-dried and ground until GSLs analysis. Five 100 mg bulks of the recombined seeds obtained in 2011 for each genotype were ground and analyzed to study the GSLs profile and GSLs content. Different agronomic traits were evaluated in the divergent selections according previous studies in this crop (Padilla et al., 2007; Vilar et al., 2008). These traits were: early vigor by using a subjective scale from 1 (very poor) to 5 (excellent); late vigor by using a subjective scale from 1 (very poor) to 5 (excellent); leaf fresh matter as the average fresh weight of a leaf (g) (mean of 25 leaves per plot taken from 5 plants per plot); leaf moisture as the percentage of fresh weight of a fresh leaf (%) and time to flowering as the number of days from transplanting until 50% of plants have the first flower.

A second assay under controlled conditions was conducted with recombined plants from the C3 (high and low) of each divergent selection and C0 in order to relate the GSL content of the plants and the expression of the principal genes related with their biosynthesis. Plants were grown in multi-pot trays in a growth chamber at 25°C ± 2°C for days and 20°C ± 2°C at night. Plants were harvested 3 months after germination and stored at −80°C until use. Three biological replicates with approximately 35 plants each one, were collected by cycle and then, each replicate were divided into three bulks. These bulks were employed to study GSL content and gene expression.

GSL Identification and Quantification

GSL extraction was conducted on samples of both trials. In field assay, GSLs were analyzed in leaves, flower buds and seeds, while in the assay under controlled conditions GSLs were analyzed in leaves. Sample extraction and desulfation, were performed according to Kliebenstein et al. (2001b) with minor modifications. Two microliters of the desulfo-GSL extract for seeds and flower buds and three microliters for leaves were used to identify and quantify the GSLs. The chromatographic analyses were carried out on an Ultra-High-Performance Liquid-Chromatograph (UHPLC Nexera LC-30AD; Shimadzu) equipped with a Nexera SIL-30AC injector and one SPD-M20A UV/VIS photodiode array detector. The UHPLC column was a C18 Atlantis® T3 waters column (3 μm particle size, 2.1 ×100 mm i.d.) protected with a C18 guard cartridge. The oven temperature was set at 30°C. Compounds were separated using the following method in aqueous acetonitrile, with a flow of 0.8 mL min^{-1}: 1.5 min at 100% H_2O, an 11 min gradient from 0% to 25% (v/v) acetonitrile, 1.5 min at 25% (v/v) acetonitrile, a minute gradient from 25% to 0% (v/v) acetonitrile, and a final 3 min at 100% H_2O. Data was recorded on a computer with the LabSolutions software (Shimadzu). All GSLs (the three major ones under selection and other minor GSLs) were

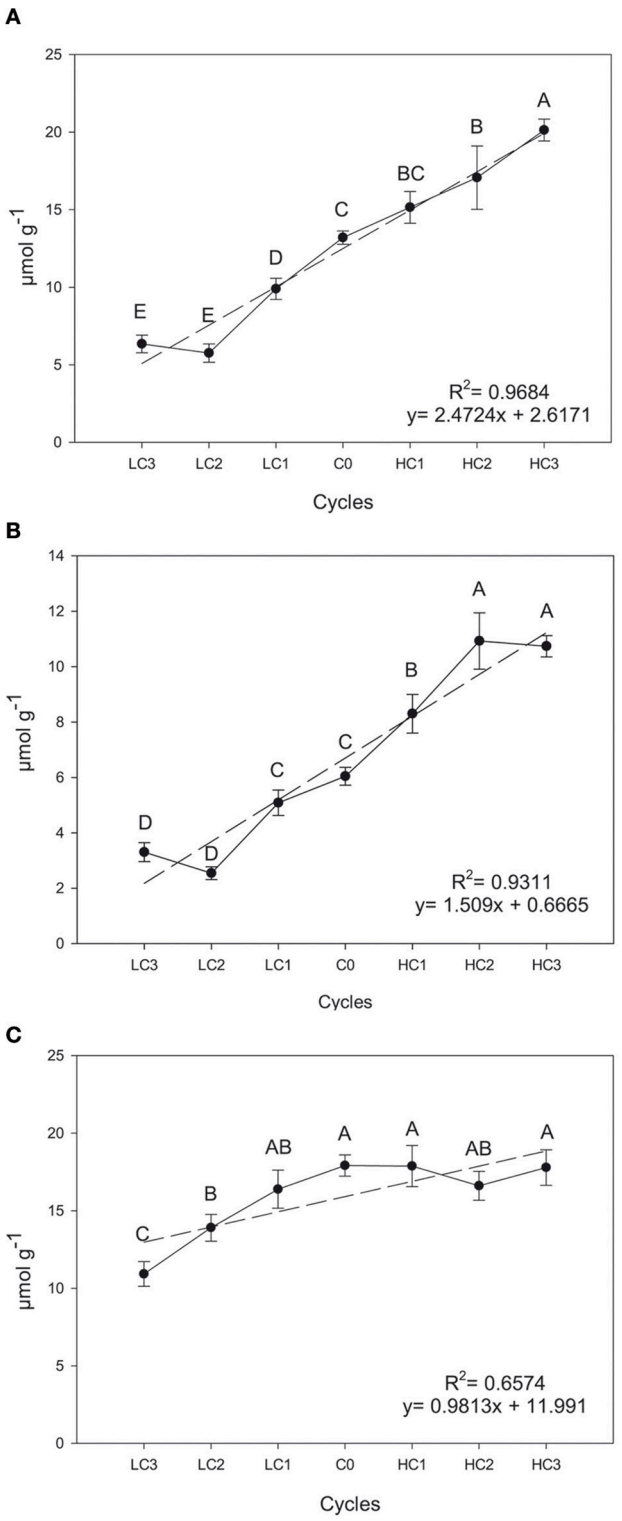


FIGURE 1 | Graphical representations of simple linear regression divergent selection in leaves for the content (μmol g^{-1}) for sinigrin (A), glucoiberin (B) and glucobrassicin (C). Data are means of three biological replicates and error bars are ± $P < 0.05$. LC1, low cycle 1; LC2, low cycle 2; LC3, low cycle 3; C0, original cycle; HC1, high cycle 1; HC2, high cycle 2; HC3, high cycle 3.

quantified at 229 nm by using SIN (sinigrin, monohydrate from Phytoplan, Diehm & Neuberger GmbH, Heidelberg, Germany) and GBS (glucobrassicin, potassium salt monohydrate, from Phytoplan, Diehm & Neuberger GmbH, Heidelberg, Germany) as external standard and expressed in μmol g^{-1} dry weight (DW). Calibration equations were made with, at least, five data points, from 0.34 to 1.7 nmol for SIN and from 0.28 to 1.4 nmol for GBS. The average regression equations for SIN, and GBS were $y = 148.818 \times (R^2 = 0.99)$, $y = 263.822 \times (R^2 = 0.99)$, respectively.

Total RNA Extraction, Primer Design, and cDNA Synthesis

Leaf RNA from three biological replicates of C0 and the C3 of each divergent selection HSINC3, LSINC3, HGIBC3, LGIBC3, HGBSC3, and LGBSC3, was isolated from 100 mg of ground samples using a Spectrum™ Plant Total RNA Kit (Quiagen, Valencia, CA, USA). Total RNA concentration was quantified using NanoDrop 1000 (Thermo Scientific, Waltham, MA, USA). To remove any traces of genomic DNA from extractions, the RNA was treated with RQ1 RNase-Free DNase (Promega, CA, USA) following the manufacturer's instructions. The cDNA was synthesized from 1 μg of total RNA using a GoScript™ Reverse Transcription System, according to the manufacturer's instructions (Promega, Madison, WI, USA).

Quantitative reverse transcription-PCR (RT-qPCR) was employed to analyze the expression patterns of 12 genes including a housekeeping gene namely glyceraldehyde-3-phosphate- dehydrogenase (GADPH) and the following genes related to the aliphatic GSLs pathway:UDP-glycosyltransferase 74B1 (UGT74B1), desulfoglucosinolate sulfotransferase (St5a), S-alkyl-thiohydroximate lyase (SUR1), glutathione S-transferase PHI 10 (GSTF10), γ-glutamyl peptidase 1 (GGP1), transcription factor (MYB51), Cytochrome P450 monooxygenase (CYP81F2) and 2-oxoglutarate-dependent dioxygenase (ALK). Finally, several genes related to the indolic and aromatic GSLs pathways were also studied: transcription factor (ATR1), cytochrome P450 83B1 (CYP83B1), Tryptophan N-monooxygenase 1(CYP79B2) and Tryptophan N-monooxygenase 2 (CYP79B3) were the aromatic and indolic genes studied.The RT-qPCR primers were designed from previously identified sequences of the GLS biosynthetic route obtained in the website http://plants.ensembl.org. Primers were designed at http://bioinfo.ut.ee/primer3-0.4.0 and they are shown in **Table S1**.

In order to determine specificity of primers designed in the current study, agarose gel electrophoresis and melting curve analyses were performed. All the primer pairs amplified single PCR products of expected size (**Table S1**) and the specificity of amplicon was confirmed by the presence of single peak during melt curve. RT-qPCR was performed using a Promega kit in a total volume of 15 μl. After denaturation at 95°C for 10 min, 40 cycles were performed under the following conditions: 95°C for 15 s and 60°C for 60 s. Primer efficiency was calculated using the LingRegPCR software (Ramakers et al., 2003) and results were normalized to GADPH expression. RT-qPCRs were carried out

TABLE 1 | Coefficients for simple linear regressions where sinigrin, glucoiberin, and glucobrassicin in leaves are the independent variables and the other GSLs present in leaves, flower buds and seeds are the dependent variables.

		SIN			GIB			GBS		
		Leaves	**Flower buds**	**Seeds**	**Leaves**	**Flower buds**	**Seeds**	**Leaves**	**Flower buds**	**Seeds**
GIB	R^2	0.2072	0.0065	0.7698		0.7102	0.5055	0.0124	0.0405	0.0876
	a	−2.636	−0.2499	−0.838**		0.6513**	0.1758**	0.1230	0.1767	−0.0920
SIN	R^2		0.5511	0.3986	0.8022	0.5466	0.3050	0.2699	0.1627	0.0187
	a		0.7688**	0.1413**	−1.044**	−0.189**	−0.078**	1.0200	0.5091	−0.0176
GBS	R^2	0.0001	0.1770	0.0037	0.8621	0.1396	0.0808		0.2873	0.2687
	a	0.0167	0.7173*	−2.7113	0.639**	0.3307	−3.4549		0.2171**	7.527**
PRO	R^2	0.5642	0.0358	0.0223	0.1589	0.0191	0.6797	0.0871	0.0473	0.0580
	a	12.325*	1.5229	0.2861*	−5.552	−0.4538	−0.773**	3.3980	2.5197	0.0774
GRA	R^2	–	0.1082	0.3042	–	0.4500	0.4772	–	0.0452	0.0294
	a	–	−8.6995	−5.9581	–	6.6729**	5.0037**	–	2.0638	1.2440
GNA	R^2	0.0928	–	0.0614	0.4870	–	0.551	0.0040	–	0.0416
	a	−19.22	–	1.4190*	−28.623	–	−1.723**	3.6650	–	0.1844
OHGBS	R^2	0.4880	0.0151	0.1161	0.0167	0.0205	0.0433	0.7135	0.2969	0.0141
	a	32.822	−3.4938	0.9019	7.937	0.3094	−0.3005	0.4861**	−8.2618**	0.1726
NeoGBS	R^2	0.0923	0.0207	0.0021	0.5209	0.0040	0	0.9331	0.6181	0.1600
	a	−4.632	0.4645	−2.9566	4.004	0.1439	0.0295	3.202**	1.0279**	7.5880
GNT	R^2	0.2049	0.0030	0.0084	0.0035	0.0299	0.4134	0.1026	0.1974	0.0556
	a	−7.821	−1.5807	−2.1207	0.773	−2.5277	−8.888**	−4.7210	3.3897*	−2.5483
Aliphatics	R^2	0.9735	0.1245	0.0203	0.1863	0.0013	0.0060	0.1442	0.0191	0.0002
	a	1.101**	0.1571	0.02234	1.000	−0.0085	−0.0061	0.368	0.0292	0.0010
Indolics	R^2	0	0.0517	0.0168	0.8599	0.0041	0.0067	0.9961	0.1270	0.0291
	a	−0.008	0.1279	0.2802	0.581**	0.0166	−0.0797	0.762**	0.0853	0.1857
Total	R^2	0.7630	0.0875	0.0201	0.9105	0.0001	0.0065	0.7341	0.0614	0.0014
	a	0.8171*	0.0725	0.0208	0.545**	0.0011	−0.0058	0.438**	0.0279	0.0025

*Aliphatic glucosinolates: GIB, Glucoiberin; SIN, Sinigrin; GRA, Glucoraphanin; GNA, Gluconapin; PRO, Progoitrin; Indolic glucosinolates: OHGBS, 4-hydroxyglucobrassicin; GBS, Glucobrassicin; NeoGBS, Neoglucobrassicin: Aromatic glucosinolate: GNT, Gluconasturtiin. R^2: coefficient of determination of each glucosinolate. a: slope of the line. *Significant at $P \leq 0.05$, and **significant at $P \leq 0.01$.*

on a 7500 Real Time PCR System (Applied Biosystem, Forster City, CA, USA).

Statistical Analysis

Combined analyses of variance across selection cycles for total and individual GSLs, agronomical traits and relative gene expression were computed using the PROC GLM of SAS v 9.2 program (SAS, 2011). Population means were compared using the Fisher protected Least Significant Difference test (LSD, $p \leq 0.05$). Besides, simple linear regression analyses were performed for the GSL implied in the three divergent selections (SIN, GIB, and GBS) as dependent variables and cycles of selection as independent variables for each organ under study (leaves, flower buds, and seeds).

Simple linear regression analyses where the GSLs under selection were the independent variables and the other GSLs, the sum of aliphatic, indolic and total GSLs were the dependent variables were also performed. Correlation coefficients between gene expression and GSLs concentration and between expressions of the different genes were computed with PROC CORR of SAS program v 9.2 (SAS, 2011).

RESULTS

Direct Response to Divergent Selection for Sinigrin, Glucoiberin, and Glucobrassicin Content in Leaves and Associated Response in Agronomical Traits

Significant and positive simple linear regression coefficients across selection cycles for SIN (R^2= 0.9684, $P \leq 0.0001$), GIB ($R^2 = 0.9311$, $P = 0.0004$) and GBS ($R^2 = 0.6574$, $P \leq 0.0001$)

concentration were observed in leaves (**Figure 1**). Generally speaking, the response to divergent selection for the three GSLs was effective and linear in leaves; therefore, mass selection is an efficient way of increasing or decreasing the concentration of individual GSLs.

The content of SIN and GIB after three cycles of divergent selection increased 52.5% ($P = 0.0074$) and 77.68% ($P = 0.0410$), respectively, and decreased 51.9% ($P = 0.0322$) and 45.33% ($P = 0.0385$), respectively. Meantime, the divergent selection performed for the leaf GBS content, was only successful and significant for decreasing the concentration, with a reduction of 39.04% ($P = 0.0248$).

Analysis of variance showed that there were not significant differences for any agronomic trait across divergent selections (data not shown).

Response to Divergent Selection for Sinigrin, Glucoiberin, and Glucobrassicin in Other Organs

There were significant and positive linear regressions between the SIN concentrations in leaves and the concentration of SIN in flower buds and seeds across selection. The same response was obtained in the other GSLs under selection, GIB and GBS although values of the R^2 for GBS were low (**Table 1**). Therefore, selection performed in leaves had a side effect in flower buds and seeds.

There were significant differences among selection cycles for the three GSLs in flower buds. Significant and positive simple linear regression coefficients for SIN ($R^2 = 0.8810$, $P = 0.0017$), GIB ($R^2 = 0.8889$, $P = 0.0015$) and GBS ($R^2 = 0.9838$, $P \leq 0.0001$) across selection cycles were found (**Figure 2**). There was a 19.7% ($P = 0.0511$) increase in SIN, a 79.62% ($P = 0.0461$) increase in GIB and a 60.02% ($P = 0.0160$) increase in GBS after three selection cycles vs. the original cycle. Meantime, the decrease in the content for SIN was 42.73% ($P = 0.0153$), 33.05% ($P = 0.0142$) for GIB and 47.60% ($P = 0.0010$) for GBS.

Positive and simple linear regressions were also found for SIN ($R^2 = 0.6889$ $P = 0.0208$), GIB ($R^2 = 0.6068$, $P = 0.0390$) and GBS ($R^2 = 0.9677$, $P = 0.0010$) in seeds (**Figure 3**). For aliphatic GSLs, selection was successful to increase the SIN and GIB concentration but selection was unsuccessful for GBS. The increase was 123.23% ($P = 0.012$) in SIN, and 661.78% ($P \leq 0.001$) in GIB and 53.35% ($P = 0.0584$) in GBS, meantime GBS was reduced in a 47.58% ($P = 0.0532$) although there are no significant differences.

Indirect Response to Divergent Selection on Other GSLs and Relationship to Gene Expression

Besides the three major GSLs under selection, this population also presents other GSLs as the aliphatics progoitrin (PRO), glucoraphanin (GRA) and gluconapin (3- butenyl, GNA), the aromatic gluconasturtiin (2-phenethyl, GNT) and the indolics, hidroxyglucobrassicin (4-hydroxy-3-indolylmethyl, OHGBS) and neoglucobrassicin (1-methoxy-3-indolylmethyl, NEOGBS) (**Table 2**). A regression analysis was made with the leaf SIN, GBS, and GIB content as independent variables and the content of the other GSLs in leaves, flower buds and seeds as dependent variables (**Table 1**). Significant and positive regressions were

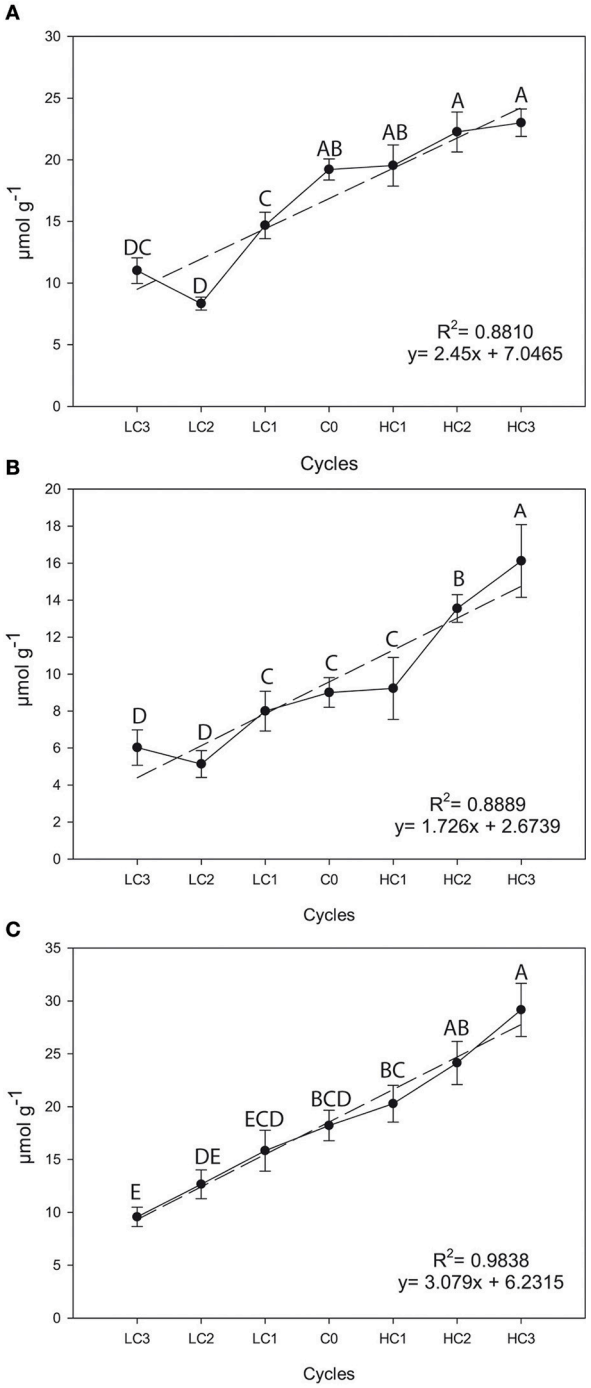


FIGURE 2 | Graphical representations of simple linear regression divergent selection in flower buds for the content (µmol g^{-1}) for sinigrin (A), glucoiberin (B) and glucobrassicin (C). Data are means of three biological replicates and error bars are ± $P < 0.05$. LC1, low cycle 1; LC2, low cycle 2; LC3, low cycle 3; C0, original cycle; HC1, high cycle 1; HC2, high cycle 2; HC3, high cycle 3.

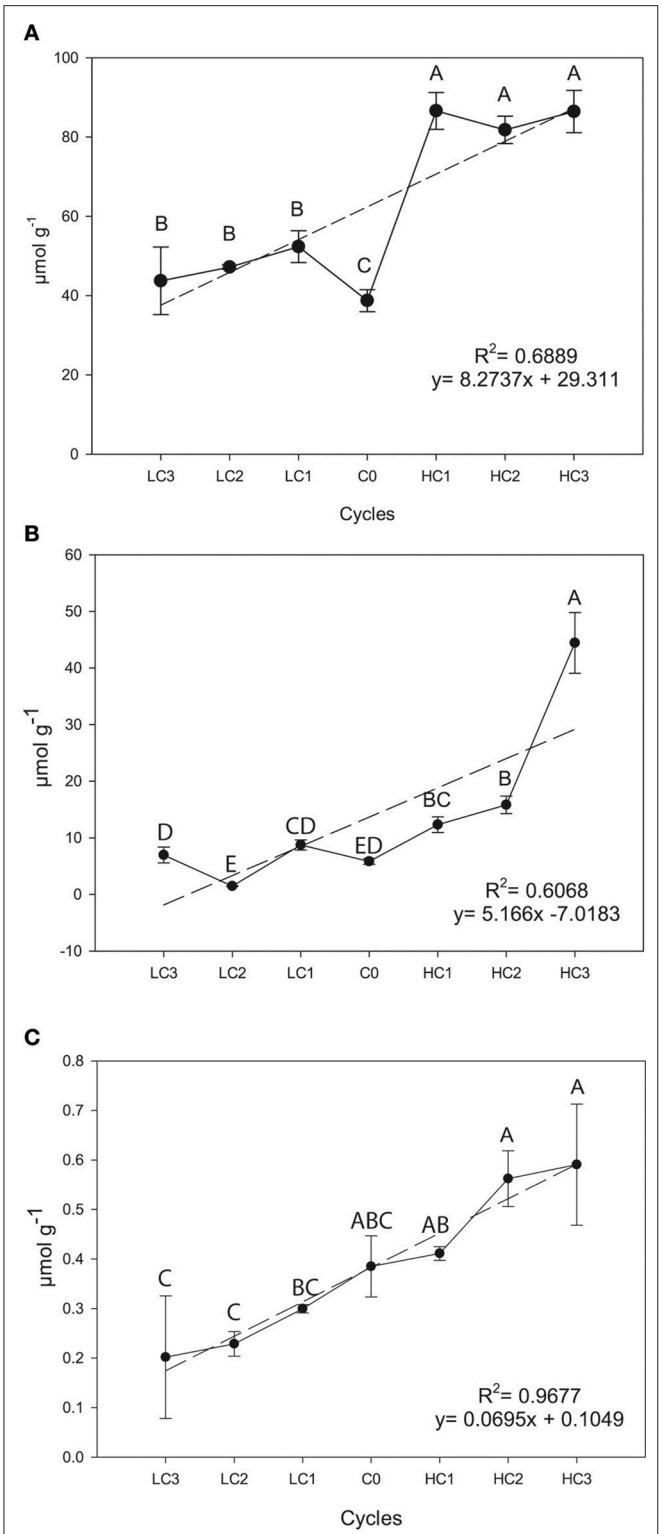


FIGURE 3 | Graphical representations of simple linear regression divergent selection in seeds for the content (µ6mol g^{-1}) for sinigrin (A), glucoiberin (B) and glucobrassicin (C). Data are means of three biological replicates, and error bars are ± $P < 0.05$. LC1, low cycle 1; LC2, low cycle 2; LC3, low cycle 3; C0, original cycle; HC1, high cycle 1; HC2, high cycle 2; HC3, high cycle 3.

TABLE 2 | Glucosinolate (GSL) concentration (µmol g^{-1} dw) of the original cycle (C0) of the kale population for the three organs under study.

Glucosinolate	Leaves[b]	Flower buds	Seeds
GIB[a]	6.045	9.002	5.831
SIN[a]	13.202	19.219	38.724
GBS[a]	17.920	18.221	0.385
PRO	1.032	1.046	8.212
GRA	0.000	0.507	0.578
GNA	0.040	0.000	2.112
OHGBS	0.250	0.418	2.127
NeoGBS	2.698	4.178	0.292
GNT	1.717	1.261	0.305
Aliphatics	20.319	29.774	56.016
Indolics	20.868	22.816	2.804
TOTAL	42.905	53.852	59.125

[a]*Glucosinolates studied in the three divergent selections.* [b]*Organ where selection was performed. Aliphatic glucosinolates: GIB, Glucoiberin; SIN, Sinigrin; GRA, Glucoraphanin; GNA, Gluconapin; PRO, Progoitrin; Indolic glucosinolates: OHGBS, 4-hydroxyglucobrassicin; GBS, Glucobrassicin; NeoGBS, Neoglucobrassicin; Aromatic glucosinolate: GNT, Gluconasturtiin.*

found between the leaf SIN content across selection cycles and PRO, aliphatic GSLs and total GSLs in leaves, GBS in flower buds and GNA in seeds. A negative correlation coefficient was found for GIB in seeds. By modifying the content of SIN, a positive related response was found in the content of PRO and GNA and a negative response in the content of GIB.

In the divergent selection program for leaf GIB content, significant and positive regressions were found between leaf GIB content and SIN and GBS, total indolic GSLs and total GSLs in leaves and GRA in flower buds and seeds (**Table 1**). Negative relationships were found between the leaf GIB content and PRO, GNA, and GNT in seeds and SIN in both seeds and flower buds.

An assay was performed to relate variation in the GSLs content in the cycles 3 (C3) of each selection with the relative expression of several genes related to their biosynthesis. The expression of 12 genes, related to the core biosynthesis of GSLs, to secondary modifications and to transcription factors were studied. The expression levels of all the genes are higher in C3 of HSIN than in C0 (**Figure 4**). Meantime in the C3 of LSIN the expression of SUR1, GSTF10, CYP79B3, CYP79B2, UGT74, MYB51, ALK, and CYP81F2 decreased respect to C0 (**Figure 4**). However, analysis of variance showed that no one of these differences are significant (data not shown), probably due to the high variability found among replicates of RT-qPCR. Relative gene expression was only significant for GGP1 in GIB selection, which expression is higher in HGIB and LGIB than in C0. However, regression analysis showed significant associations between gene expression and GSLs content across selection cycles.

High and significant correlations between SIN and GIB concentration and GSL-ALK expression were found (**Table 3**, **Figure 4**). When SIN concentration increases, GSL-ALK gene expression also increases ($r = 0.92$); however, when GIB concentration increases, the expression of GSL-ALK gene decreases ($r = -0.82$). Variation in SIN concentration causes

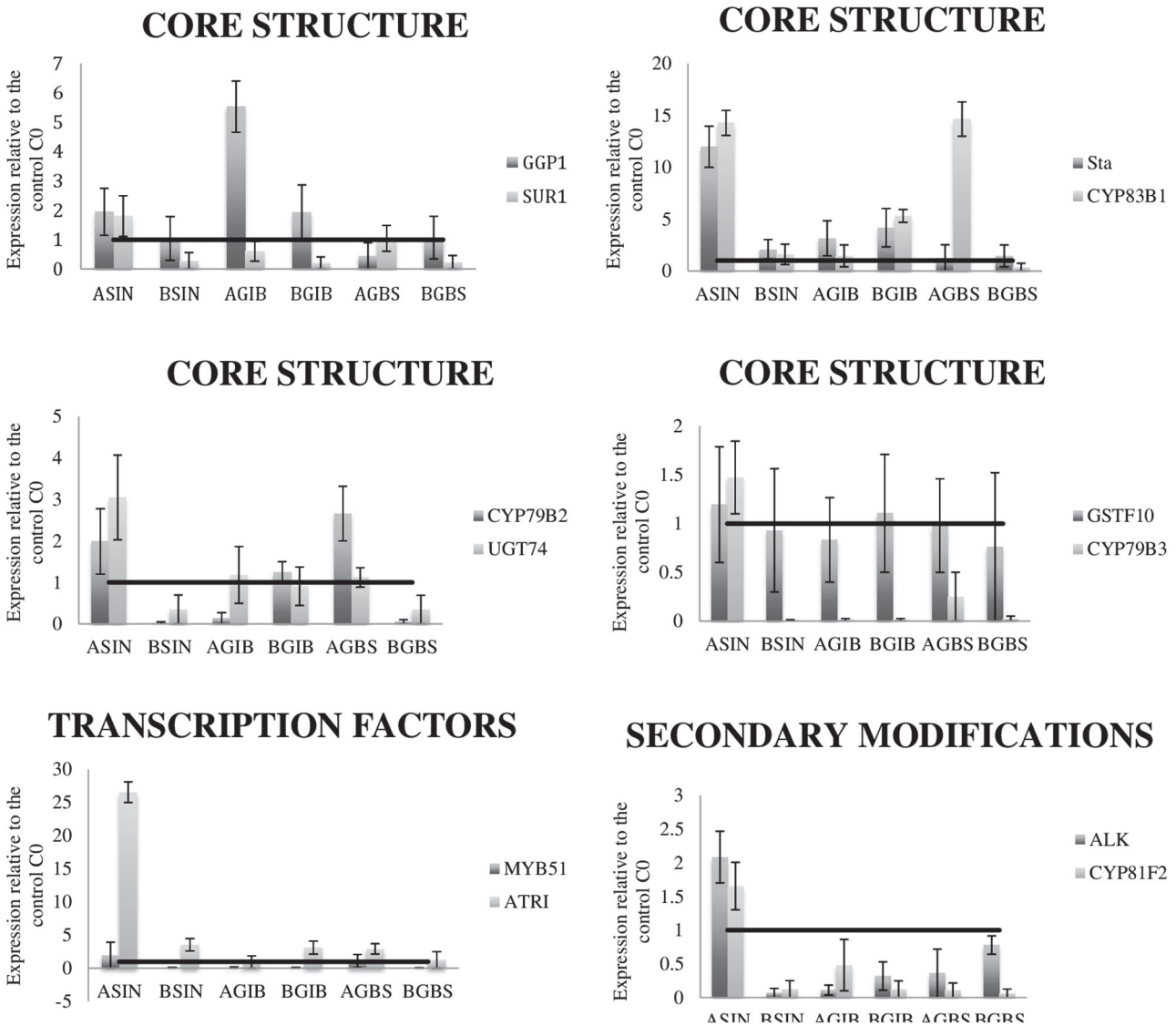


FIGURE 4 | Graphical representation of gene expression levels in the C3 of the three divergent selections relative to the expression of the same genes in the control (C0). Data are means of three biological replicates, and the error bars indicate their standard deviations. The horizontal line in each graph indicates the relative expression of each gene in the control. LC1, low cycle 1; LC2, low cycle 2; LC3, low cycle 3; C0, original cycle; HC1, high cycle 1; HC2, high cycle 2; HC3, high cycle 3.

significant and positive correlations with the majority of the other genes studied while variation in GIB only causes a correlated and negative response in CYP79B3 (**Table 3**, **Figure 4**). In the case of SIN selection, expression of GSL-ALK gene presents a positive and significant correlation with CYP79B2 ($r = 0.967$), CYP79B3 ($r = 0.991$), CYP83B1 ($r = 0.958$), SUR1 ($r = 0.958$), UGT74B1 ($r = 0.971$) and St5a ($r = 0.942$). Meantime, in the case of GIB divergent selection, GSL-ALK expression showed significant and positive correlations with CYP79B2 ($r = 0.780$) and CYP83B1 ($r = 0.966$) genes.

In the divergent selection for the leaf GBS content, significant and positive regression was found with the content of OHGBS, NEOGBS, total indolic GSLs and total GSLs (**Table 1**). GBS is the precursor of OHGBS and NeoGBS in the biosynthetic pathway of indolic GSLs (**Figure S2B**); therefore, variation in GBS content provokes a positive response in the leaf content of NeoGBS and OHGBS.

As in the case of aliphatic selections, there were no significant differences in gene expression across selection cycles. However, significant correlations of gene expression with GBS content were found. The expression of CYP83B1 was positively correlated to GBS content (**Figure 4**). This gene is responsible of the conversion of GBS into OHGBS.

A significant regression of GBS with the aromatic GSL GNT was found in flower buds although the R^2 was low. Indolic and aromatic GSLs share the gene UGT74B1 in their pathways

TABLE 3 | Significant correlations between Sinigrin (SIN), Glucoiberin (GIB) and Glucobrassicin (GBS) concentration and expression of 12 genes related to the glucosinolate biosynthetic route.

		Core structure								Secondary modifications	
		CYP79B2	CYP79B3	CYP83B1	GSTF10	GGP1	SUR1	UGT74B1	St5a	CYP81F2	ALK
SIN	Correlation	0.8356	0.9377	0.9538	0.0116	−0.5030	0.8171	0.8712	0.8837	0.8389	0.9206
	P-value	0.0103	0.0015	0.0012	0.1052	0.0770	0.0083	0.0076	0.0008	0.0057	0.0012
GIB	Correlation	−0.2179	−0.9475	−0.2699	0.0381	0.6680	0.5567	−0.4150	0.3552	0.2622	−0.8276
	P-value	0.0678	0.0001	0.0756	0.0666	0.0599	0.0601	0.0764	0.1222	0.1323	0.0033
GBS	Correlation	−0.1829	−0.1159	0.2985	0.1302	0.0249	0.4752	0.8485	0.4935	0.7879	−0.5388
	P-value	0.1211	0.0644	0.0652	0.0989	0.0657	0.0654	0.0019	0.0696	0.0431	0.0635

(**Figure S2B**), which expression was modified with the content of GBS (**Figure 4**).

The relationship of the expression of the indolic regulators, ATRI and MYB51 genes with the content of GBS was also studied, but they were no significant.

DISCUSSION

Direct Response to Divergent Selection for Sinigrin, Glucoiberin, and Glucobrassicin Content in Leaves and Associated Response in Agronomical Traits

After three cycles of divergent selection, the response to divergent selection for the three GSLs under study was effective and linear in kale leaves. The effect of selecting for GSLs content did not have any effect in agronomical traits; therefore, mass selection is an efficient way of increasing or decreasing the concentration of individual GSLs. A modification in the concentration of the aliphatic GSLs (SIN and GIB) was observed in both senses of the divergent selections. Stowe and Marquis (2011) obtained similar results in a divergent selection to modify the content of total GSLs in *B. rapa*. However, the divergent selection performed for the leaf GBS content was only successful and significant for decreasing the concentration. The asymmetric response in a divergent selection program, such as we found for GBS content, has been found before, for example in maize for leaf chlorophyll content but the cause is still unknown (Korkovelos and Goulas, 2011). There are some possible causes to explain this effect such as differential selection, genetic asymmetry, selection for heterozygotes, inbreeding depression or maternal effects (Falconer, 1989).

The mass selection is an effective method for highly heritable traits. Although the estimates of heritability could not be calculated with the experimental design used in our work, according to the results obtained, we can conclude that heritability should be high enough. In this sense, Madsen et al. (2014) in *B. napus* and Márquez-Lema et al. (2009) in *B. carinata,* estimated the heritability for total GSLs in seeds with values of $h^2 = 0.90$ and $h^2 = 0.58$, respectively. In another study, Van Doorn et al. (1998) established the heritability for two aliphatic GSLs (SIN and PRO) in different cultivars of Brussels sprouts with values of $h^2 = 0.77$ and $h^2 = 0.79$, respectively.

Response to Divergent Selection for Sinigrin, Glucoiberin, and Glucobrassicin in Other Organs

Leaves are the organ most consumed in kales, hence the importance to perform the divergent selections for specific GSLs in this organ. It has long been known that also there are GSLs in other organs such as roots, shoots, stems or seeds (Grubb and Abel, 2006) in part by the new GSLs biosynthesis or by translocation from other organs. We hypothesized that GSLs content on other organs, such as flower buds and seeds, could be affected by the selections performed in leaves.

When selection is carried out to increase the content of the three GSLs in leaves, there is also an increase of the same GSLs in flower buds and seeds except for SIN in flower buds and GBS in seeds. When the selection is carried out for decreasing the content of the three GSLs in leaves, there is also a reduction of the same GSLs in flower buds and no related responses were found in seeds for both aliphatic and indolic GSLs. The reproductive organs, including seeds, flowers and fruits, which contribute most to plant fitness, are expected to have the highest concentrations of GSLs. In this way, Brown et al. (2003) in *A. thaliana* demonstrate that seeds present higher content of GSLs than vegetative organs. GSLs accumulation represents the net effect of biosynthesis, transport and catabolism. It can be possible that, by modifying the action of genes responsible for the concentration of GSLs in leaves, the action of the same genes were also modified in flower buds and seeds.

Differences in concentration and pattern of GSLs in different organs of *B. rapa* were related to differential expression of transcription factors involved in GSLs biosynthesis (Clarke, 2010). In our case, the same response was found in leaves, flower buds and seeds; therefore, genes related to biosynthetic pathway and no transcription factors could be implied in the divergent selection. Besides, there is a translocation of GSLs from vegetative organs to reproductive ones with the development. Du and Halkier (1998) observed that the high accumulation of GSLs in seeds is not connected with a corresponding high

level of associated biosynthesis, suggesting the involvement of transport processes. Chen et al. (2001) demonstrated the translocation of radiolabeled p-hydroxybenzyl GSL from leaves to seeds via phloem, either exogenously applied or de novo synthesized. In fact, a recent study in *A. thaliana* shows the necessary presence of one specific transporter for the GSL translocation from other organs to seeds (Nour-Eldin et al., 2012) and the necessary presence of these transporters related with the movement of GSLs from roots to shoots (Madsen et al., 2014).

Indirect Response to Divergent Selection on Other GSLs and Relationship to Gene Expression

The kale population studied in this work presents other GSLs as the aliphatics PRO, GRA, and GNA, the aromatic GNT and the indolic GSL, OHGBS, and NEOGBS which content could have been modified indirectly by the divergent selection performed on leaves for SIN, GIB, and GBS.

In the divergent selection for SIN, a positive correlation was found with the content of PRO and GNA and a negative correlation with the content of GIB suggesting that modification in the SIN concentration by selection is related to the GSL-ALK locus. GSL profile in Brassicaceae can be partially explained by genetic variation in the GSL-ALK locus encoding (2-oxoglutarate-dependent dioxygenase) which catalyzes the conversion of methylsulfinylalkyl GSL to the alkenyl form in plants (Li and Quiros, 2003). In the biosynthetic pathway of GSLs, the locus GSL-ALK controls the side chain desaturation and its presence determines the production of the alkenyl GSLs SIN (3C-GSL) and PRO and GNA (4C-GSL) (Li et al., 2001) (**Figure S2A**).

In the divergent selection program for leaf GIB content, there is a negative correlation of leaf content of GIB on the content of SIN, PRO and GNA, and a positive correlation on the content of GRA, which suggests that the modification of the content of GIB is related to the major locus, GSL-ALK. In the biosynthetic pathway of aliphatic 3C-GSLs, the alkenization of GIB produces SIN. In the pathway of 4C-GSLs, the alkenization of GRA produces GNA, which is afterwards transformed into PRO. Alkenizations are carried out by the GSL-ALK locus.

Supporting the role of GSL-ALK in modifying the content of SIN and GIB, high and significant correlations between SIN and GIB concentration and GSL-ALK expression were found. When SIN concentration increases, GSL-ALK gene expression also increases; however, when GIB concentration increases, the expression of GSL-ALK gene decreases. These results showed that different alleles of the GSL-ALK might be implied in these selections. The expression of GSL-ALK is correlated with the expression of as the genes CYP79B2, CYP79B3, CYP83B1, SUR1, UGT74B1, and St5a, all of them related to the synthesis of the core structure of aliphatic GSLs.

Recent evidence suggests a potential for feedback regulation in the GSL pathway. Genetic variation at GSL-ALK locus is linked to the production of alkenyl GSLs, but also to increase of total aliphatic GSL in *A. thaliana* (Kliebenstein et al., 2001a; Wentzell et al., 2007). Expression of the homologous of GSL-ALK (AOP2) from *B. oleracea* in a naturally occurring knockout genotype of *A. thaliana*, lead to the accumulation of alkenyl GSLs, doubling of total aliphatic GSL content and the induction of aliphatic GSL biosynthetic and regulatory genes. Wentzell et al. (2007) proposed that GSL-ALK has a regulatory effect in other genes of GSL synthesis trough a mechanism that is still unknown. More recently, Sotelo et al. (2014) found that GSL-ALK plays a central role in a network of epistatic interactions between ten QTLs related to GSLs, suggesting a possible regulatory effect of this locus in the GSL pathway.

By modifying the content of SIN, a positive response is also found for GBS and total indolic GSLs. Sotelo et al. (2014) found that GSL-ALK controls indirectly the variability for GBS content by epistatic interactions, indicating a cross talk between indolic and aliphatic pathways.

In the divergent selection for the leaf GBS content, results showed a significant and positive relationship with the content of OHGBS, NEOGBS, total indolic GSLs and total GSLs. GBS is the precursor of OHGBS and NeoGBS in the biosynthetic pathway of indolic GSLs (**Figure S2B**); therefore, variation in GBS content provokes a positive response in the leaf content of NeoGBS and OHGBS. In this case, we only found significant coefficients in leaves and flower buds, probably because the GBS levels in seeds are too low. Confirming these results, GBS content was correlated with CYP81F2 gene expression (**Table 3**), which catalyzes the conversion of GBS to OHGBS (Pfalz et al., 2009).

The relationship between GBS content and GSL GNT was only detected in flower buds, probably due to the higher concentration of GNT in flower buds than in leaves or seeds. Ours results suggest that expression UGT74B1 gene that is involved in the indolic and aromatic GSLs pathway (**Figure S2B**) was modified in relation with the content of GBS.

CONCLUSIONS

Divergent mass selection for SIN, GIB, and GBS leaf content was successful indicating that there is high genetic variability within the population which allows us to modify the concentration of GSLs through mass selection. The genotypes obtained in this study (with increased and decreased GSL content) can represent valuable materials for undertaking basic studies about the biological effects of the major GSLs present in kales.

There was a side effect of divergent selection performed in leaves in the GSL content of flower buds and seeds, indicating modification of the synthesis of GSLs in these organs or translocation of GSLs from leaves. A further study to examine GSL-related gene expression changes, particularly GSL-ALK, in seeds and flower buds would be necessary to conclude if the changes of GSL contents in leaves during selection were caused by the reallocation of GSLs among different tissues/organs within plant or changes of GSL-related gene expression in leaves or both. Because kale plants have long vegetative periods (they are biannual), large heights, and it is very difficult to grow them in

culture chambers to obtain flower buds, a new experiment in the field would be required in order to collect the buds and the seeds of each divergent selection and to perform further gene expression analyses.

Indirect effects of divergent selection performed for the two aliphatic GLS under selection (SIN and GIB) in the content of other GSLs suggest that different alleles of the locus GSL-ALK are responsible for the variation across the selection cycles. The expression of genes involved in the GSLs pathway confirmed these results. At the same time, this locus could be responsible of the indirect response found for the indolic GBS. In the case of indolic divergent selection, CYP81F2 gene could be the responsible of the variations in concentration across the selection cycles.

AUTHOR CONTRIBUTIONS

TS carried out the experiments and wrote the manuscript. TS, PS, and VR performed the genetic analysis. TS and PV performed the glucosinolate analysis. PV, MC, and PS conceived the study and participated in its design. MC and PV coordinated the work. All authors have read and approved the manuscript.

FUNDING

This work was supported by the National Plan for Research and Development AGL-2012-35539, AGL2015-66256-C2-1-R and financed by the European Regional Development Funds (FEDER).

ACKNOWLEDGMENTS

The authors thanks to Rosaura Abilleira, César González, and Pilar Comesaña for laboratory help and field work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01012

Figure S1 | Genetic model for the two first steps of the glucosinolates synthesis. Extracted and modified of Redovnikovic et al. (2008).

Figure S2 | A biochemical genetic model of the biosynthesis of aliphatic glucosinolates (A) and indolic glucosinolates (B) in *Brassicaceae* including the major genes controlling this process.

Table S1 | Primer sequences used for RT-qPCR and gene expression analysis.

REFERENCES

Bak, S., Feyereisen, R., Tax, F. E., Feldmann, K. A., and Galbraith, D. W. (2001). CYP83B1, a cytochrome P450 at the metabolic branch point in auxin and indole glucosinolate biosynthesis in *Arabidopsis*. *Plant Cell* 13, 101–111. doi: 10.1105/tpc.13.1.101

Bekaert, M., Edger, P., Hudson, C., Pires, C., and Conant, G. (2012). Metabolic and evolutionary costs of herbivory defense: systems biology of glucosinolate synthesis. *New Phytol.* 196, 596–605. doi: 10.1111/j.1469-8137.2012.04302.x

Brown, P., Tokuhisa, J., Reichelt, M., and Gershenzon, J. (2003). Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* 62, 471–481. doi: 10.1016/S0031-9422(02)00549-6

Cartea, M. E., and Velasco, P. (2008). Glucosinolates in brassica foods: bioavailability in food and significance for human health. *Phytochem. Rev.* 7, 213–229. doi: 10.1007/s11101-007-9072-2

Chen, S. X., Petersen, B. L., Olsen, C. E., Schulz, A., and Halkier, B. A. (2001). Long-distance phloem transport of glucosinolates in *Arabidopsis*. *Plant Physiol.* 127, 194–201. doi: 10.1104/pp.127.1.194

Clarke, D. (2010). Glucosinolates, structures and analysis in food. *Anal. Methods* 2, 310–325. doi: 10.1039/b9ay00280d

Du, L. C., and Halkier, B. A. (1998). Biosynthesis of glucosinolates in the developing silique walls and seeds of *Sinapis alba*. *Phytochemistry* 48, 1145–1150. doi: 10.1016/S0031-9422(97)00877-7

Fahey, J. W., and Stephenson, K. K. (1999). Cancer chemoprotective effects of cruciferous vegetables. *HortScience* 34, 1159–1163.

Fahey, J. W., Talalay, P., and Zalcmann, A. T. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51. doi: 10.1016/S0031-9422(00)00316-2

Falconer, D. S. (1989). "Selección:II. Resultados de los experimentos," in *Introducción a la Genética Cuantitativa*, ed Longman group limited (Mexico: CECSA), 199–218.

Forte, A., De Sanctis, R., Leonetti, G., Manfredelli, S., and Urbano, V. (2008). Dietary chemoprevention of colorectal cancer. *Ann. Ital. Chir.* 79, 261–267.

Grubb, C. D., and Abel, S. (2006). Glucosinolate metabolism and its control. *Trends Plant Sci.* 11, 89–100. doi: 10.1016/j.tplants.2005.12.006

Halkier, B. A., and Du, L. C. (1997). The biosynthesis of glucosinolates. *Trends Plant Sci.* 2, 425–431. doi: 10.1016/S1360-1385(97)90026-1

Kliebenstein, D. J., Kroymann, J., Brown, P., Figuth, A., and Pedersen, D. (2001a). Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol.* 126, 811–825.

Kliebenstein, D. J., Lambrix, V., Reichelt, M., Gershenzon, J., and Mitchell-Olds, T. (2001b). Gene duplication and the diversification of secondary metabolism: side chain modification of glucosinolates in *Arabidopsis thaliana*. *Plant Cell* 13, 681–693. doi: 10.1105/tpc.13.3.681

Korkovelos, A. E., and Goulas, C. K. (2011). Divergent mass selection for leaf chlorophyll content measured using chlorophyll meter readings in a maize composite population. *Crop Sci.* 51, 1437–1443. doi: 10.2135/cropsci2010.04.0239

Kushad, M., Brown, A., Kurilich, A., Juvik, J., Klein, B., Wallig, M., et al. (1999). Variation of glucosinolates in vegetable crops of *Brassica oleracea*. *J. Agric. Food Chem.* 47, 1541–1548. doi: 10.1021/jf980985s

Li, G., and Quiros, C. F. (2003). In planta side-chain glucosinolate modification in *Arabidopsis* by introduction of dioxygenase *Brassica* homolog BoGSL-ALK. *Theor. Appl. Genet.* 162, 1937–1943. doi: 10.1007/s00122-002-1161-4

Li, G., Riaz, A., Goyal, S., Abel, S., and Quiros, C. F. (2001). Inheritance of three major genes involved in the synthesis of aliphatic glucosinolates in *Brassica oleracea*. *J. Am. Soc. Horticult. Sci.* 126, 427–431.

Liu, Z., Hirani, A., McVetty, P. B. E., Daayf, F., Quiros, C. F., and Li, G. (2012). Reducing progoitrin and enriching glucoraphanin in *Brassica napus* seeds through silencing of the GSL-ALK gene family. *Plant Mol. Biol.* 79, 179–189. doi: 10.1007/s11103-012-9905-2

Madsen, S. R., Olsen, C. E., Nour-Eldin, H. H., and Halkier, B. A. (2014). Elucidating the role of transport processes in leaf glucosinolate distribution. *Plant Physiol.* 166, 1450–1462. doi: 10.1104/pp.114.246249

Márquez-Lema, A., Fernandez-Martinez, J. M., Perez-Vich, B., and Velasco, L. (2009). Inheritance of very high glucosinolate content in Ethiopian mustard seeds. *Plant Breed.* 128, 278–281. doi: 10.1111/j.1439-0523.2008.01563.x

Mikkelsen, M. D., Petersen, B. L., Olsen, C. E., and Halkier, B. A. (2002). Biosynthesis and metabolic engineering of glucosinolates. *Amino Acids* 22, 279–295. doi: 10.1007/s007260200014

Mikkelsen, M. D., Wittstock, U., Hansen, C. H., and Halkier, B. A. (2000). Cytochrome P450 CYP79B2 from *Arabidopsis* catalyzes the conversion of tryptophan to indole-3-acetaldoxime, a precursor of indole glucosinolates and indole-3-acetic acid. *J. Biol. Chem.* 275, 33712–33717. doi: 10.1074/jbc.M001667200

Mithen, R., Fulkner, K., Magrath, R., Rose, P., Williamson, G., and Marquez, J. (2003). Development of isothiocyanate-enriched broccoli, and its enhanced ability to induce phase 2 detoxification enzymes in mammalian cells. *Theor. Appl. Genet.* 106, 727–734. doi: 10.1007/s00122-002-1123-x

Naur, P., Petersen, B., Mikkelsen, M., Bak, S., Rasmussen, H., Olsen, C., et al. (2003). CYP83A1 and CYP83B1, two nonredundant cytochrome P450 enzymes metabolizing oximes in the biosynthesis of glucosinolates in *Arabidopsis*. *Plant Physiol.* 133, 63–72. doi: 10.1104/pp.102.019240

Nour-Eldin, H. H., Andersen, T. G., Burow, M., Madsen, S. R., Jorgensen, M. E., Olsen, C. E., et al. (2012). NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* 488, 531–534. doi: 10.1038/nature11285

Padilla, G., Cartea, M. E., and Ordás, A. (2007). Comparison of several clustering methods in grouping kale landraces. *J. Am. Soc. Horticult. Sci.* 132, 387–395.

Pfalz, M., Vogel, H., and Kroymann, J. (2009). The gene controlling the indole glucosinolate modifier1 quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis*. *Plant Cell* 21, 985–999. doi: 10.1105/tpc.108.063115

Ramakers, C., Rijter, J. M., Lekanne-Deprez, R. H., and Moorman, A. F. M. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.* 13, 62–66. doi: 10.1016/S0304-3940(02)01423-4

Redovnikovic, I., Glivetic, T., Delonga, K., and Vorkapic, F. (2008). Glucosinolates and their potential role in plant. *Period. Biol.* 110, 297–309.

Röbbelen, G., and Thies, W. (1980). "Biosynthesis of seed oil and breeding for improved oil quality of rapeseed," in *Brassica Crops and Wild Allies*, *Biology and Breeding*, eds S. Tsunoda, K. Hita, and C. Gomez-Campo (Tokio: Japan Scientific Societies Press), 253–283.

Santolamazza–Carbone, S., Velasco, P., Soengas, P., and Cartea, M. E. (2014). Bottom-up and top-down herbivore regulation mediated by glucosinolates in *Brassica oleracea* var. *acephala*. *Oecologia* 174, 893–907. doi: 10.1007/s00442-013-2817-2

Sarikamis, G., Marquez, J., MacCormack, R., Bennett, R., and Roberts, J. (2006). High glucosinolate broccoli: a delivery system for sulforaphane. *Mol. Breed.* 18, 219–228. doi: 10.1007/s11032-006-9029-y

SAS (2011). *Institute Inc SAS® 9.2 SAS. Cary, NC*: Enhanced Logging Facilities, SAS Institute.

Sorensen, H. (1988). Glucosinolates-structure, properties, function. *Abstracts of Papers in American Chemical Society 195*, 79-AGFD.

Sotelo, T., Soengas, P., Velasco, P., Rodriguez, V. M., and Cartea, M. E. (2014). Identification of metabolic QTLs and candidate genes for glucosinolate synthesis in *Brassica oleracea* leaves, seeds and flower buds. *PLoS ONE* 9:e91428. doi: 10.1371/journal.pone.0091428

Stefansson, B. R., and Kondra, Z. P. (1975). Tower summer rape. *Can. J. Plant Sci.* 55, 343–344. doi: 10.4141/cjps75-053

Stowe, K., and Marquis, R. (2011). Costs of defense: correlated responses to divergent selection for foliar glucosinolate content in *Brassica rapa*. *Evol. Ecol.* 25, 763–775. doi: 10.1007/s10682-010-9443-9

Van Doorn, H. E., Van Der Kruk, G. C., Van Holst, G. J., Raaijmakers-Ruijs, N. C. M. E., Postma, E., Groeneweg, B., et al. (1998). The glucosinolates sinigrin and progoitrin are important determinants for taste preference and bitterness of brussels sprouts. *J. Sci. Food Agric.* 78, 30–38.

Van Horn, L., McCoin, M., Kris Etherton, P., and Burke, F. (2008). The evidence for dietary prevention and treatment of cardiovascular disease. *J. Am. Diet. Assoc.* 108, 287–331. doi: 10.1016/j.jada.2007.10.050

Velasco, P., Cartea, M. E., Gonzalez, C., Vilar, M., and Ordás, A. (2007). Factors affecting the glucosinolate content of kale (*Brassica oleracea acephala* group). *J. Agric. Food Chem.* 55, 955–962. doi: 10.1021/jf0624897

Vilar, M., Cartea, M. E., Padilla, G., Soengas, P., and Velasco, P. (2008).The potential of kales as a promising vegetable crop. *Euphytica* 159, 153–165 doi: 10.1007/s10681-007-9468-5

Virgili, F., and Marino, M. (2008). Regulation of cellular signals from nutritional molecules: a specific role for phytochemicals, beyond antioxidant activity. *Free Radic. Biol. Med.* 45, 1205–1216. doi: 10.1016/j.freeradbiomed.2008.08.001

Wentzell, A., Rowe, H., Hansen, B., Ticconi, C., Halkier, B., and Kliebenstein, D. (2007). Linking metabolic QTls with network and cis e-QTLs controlling biosynthetic pathways. *PLoS Genet.* 3:e162. doi: 10.1371/journal.pgen.0030162

Zukalova, H., and Vasak, J. (2002). The role and effects of glucosinolates of *Brassica* species - a review. *Rostlinna Vyroba* 48, 175–180.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Sotelo, Velasco, Soengas, Rodríguez and Cartea. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 February 2016
doi: 10.3389/fpls.2016.00045

Plants as Biofactories: Postharvest Stress-Induced Accumulation of Phenolic Compounds and Glucosinolates in Broccoli Subjected to Wounding Stress and Exogenous Phytohormones

*Daniel Villarreal-García[1], Vimal Nair[2], Luis Cisneros-Zevallos[2] and Daniel A. Jacobo-Velázquez[1]**

[1] Centro de Biotecnología-FEMSA, Tecnológico de Monterrey – Campus Monterrey, Monterrey, Mexico, [2] Department of Horticultural Sciences, Texas A&M University, College Station, TX, USA

OPEN ACCESS

Edited by:
Narendra Tuteja,
International Centre for Genetic Engineering and Biotechnology, India

Reviewed by:
Inger Martinussen,
NIBIO Norwegian Institute for Bioeconomy, Norway
Laura Jaakola,
UiT The Arctic University of Norway, Norway

***Correspondence:**
Daniel A. Jacobo-Velázquez
djacobov@itesm.mx

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: 04 October 2015
Accepted: 11 January 2016
Published: 10 February 2016

Citation:
Villarreal-García D, Nair V, Cisneros-Zevallos L and Jacobo-Velázquez DA (2016) Plants as Biofactories: Postharvest Stress-Induced Accumulation of Phenolic Compounds and Glucosinolates in Broccoli Subjected to Wounding Stress and Exogenous Phytohormones. Front. Plant Sci. 7:45. doi: 10.3389/fpls.2016.00045

Broccoli contains high levels of bioactive molecules and is considered a functional food. In this study, postharvest treatments to enhance the concentration of glucosinolates and phenolic compounds were evaluated. Broccoli whole heads were wounded to obtain florets and wounded florets (florets cut into four even pieces) and stored for 24 h at 20 °C with or without exogenous ethylene (ET, 1000 ppm) or methyl jasmonate (MeJA, 250 ppm). Whole heads were used as a control for wounding treatments. Regarding glucosinolate accumulation, ET selectively induced the 4-hydroxylation of glucobrassicin in whole heads, resulting in ∼223% higher 4-hydroxyglucobrassicin than time 0 h samples. Additionally, glucoraphanin was increased by ∼53% in whole heads treated with ET, while neoglucobrassicin was greatly accumulated in wounded florets treated with ET or MeJA, showing increases of ∼193 and ∼286%, respectively. On the other hand, although only whole heads stored without phytohormones showed higher concentrations of phenolic compounds, which was reflected in ∼33, ∼30, and ∼46% higher levels of 1,2,2-trisinapoylgentiobose, 1,2-diferulolylgentiobiose, and 1,2-disinapoyl-2-ferulolylgentiobiose, respectively; broccoli florets stored under air control conditions showed enhanced concentrations of 3-*O*-caffeoylquinic acid, 1,2-disinapoylgentiobiose, and 1,2-disinapoyl-2-ferulolylgentiobiose (∼22, ∼185, and ∼65% more, respectively). Furthermore, exogenous ET and MeJA impeded individual phenolics accumulation. Results allowed the elucidation of simple and effective postharvest treatment to enhance the content of individual glucosinolates and phenolic compounds in broccoli. The stressed-broccoli tissue could be subjected to downstream processing in order to extract and purify bioactive molecules with applications in the dietary supplements, agrochemical and cosmetics markets.

Keywords: broccoli, glucosinolates, phenolic compounds, wounding stress, ethylene, methyl jasmonate, neoglucobrassicin

INTRODUCTION

Broccoli (*Brassica oleracea* L. var. *Italica*) is a very important crop in economic terms. According to the Food and Agricultural Organization of the United Nations statistical database (FAOSTAT), in the year 2013 ∼22 million tons of broccoli and cauliflowers were produced worldwide. Broccoli production and consumption per capita has greatly increased over the last two decades. From 1993, broccoli worldwide production augmented by ∼120% (Food and Agricultural Organization of the United Nations [FAOSTAT], n.d.), whereas broccoli consumption per capita increased by ∼50% in the United States (Economics, Statistics, and Market Information System [ERS], n.d.). The increased economic importance of broccoli is in part due to an increase in the number of consumers interested in eating more functional foods (Agricultural Marketing Resource Center [AgMRC], n.d.).

Broccoli contains high levels of phenolics and glucosinolates, which are among the most effective bioactive molecules that prevent chronic and degenerative diseases (Vinson et al., 1998; Fahey and Talalay, 1999; Heo et al., 2007; Rodríguez-Cantú et al., 2011). Phenolic compounds are widely known as potent antioxidants (Rice-Evans et al., 1997; Jacobo-Velázquez and Cisneros-Zevallos, 2009; Brewer, 2011; Del Rio et al., 2013). Likewise, glucosinolates are amino acid-derived secondary metabolites that when hydrolyzed by a β-thioglucosidase (myrosinase) yield isothiocyanates (Radojčić-Redovniković et al., 2008), which are strong inducers of phase II enzymes, helping to prevent oxidative stress caused by reactive electrophile species (Fahey and Talalay, 1999). Besides, sulforaphane, one of the most common isothiocyanates in broccoli, has been shown to eradicate infections by *Helicobacter pylori* (Fahey et al., 2002) and inhibit chronic inflammatory processes (Juurlink, 2001).

It has been reported that the application of postharvest abiotic stresses (i.e., wounding, UV-light radiation, and exogenous phytohormones) induce the accumulation of health-promoting compounds in plants (Cisneros-Zevallos, 2003). In addition, when horticultural crops are subjected to extreme postharvest abiotic stress conditions, the genetic potential of plants to produce secondary metabolites can be exploited inducing the accumulation of high levels of bioactive molecules (Jacobo-Velázquez and Cisneros-Zevallos, 2012). In the particular case of broccoli, the application of certain postharvest abiotic stress conditions could be used as an approach to induce the activation of phenolic and glucosinolate biosynthesis pathways, leading to an enhancement of its nutraceutical content. This becomes particularly relevant when alternative uses for horticultural crops are needed; especially in fresh produce not meeting quality standards for human consumption, which represent one third of worldwide production (Food and Agriculture Organization of the United Nations [FAO], n.d.).

Plant hormones such as methyl jasmonate (MeJA) and ethylene (ET) have been used as elicitors of high-value antioxidants in several plant models. MeJA, a phytohormone involved in diverse developmental processes and plant defense mechanisms, has been studied as a pre-harvest elicitor to enrich glucosinolate and/or phenolic content in *Brassica rapa* and broccoli (Liang et al., 2006; Kim and Juvik, 2011; Ku et al., 2014; Liu et al., 2014). Moreover, ET has also been shown to be effective in the activation of glucosinolate biosynthetic genes in *Arabidopsis* (Mikkelsen et al., 2003). On the other hand, the accumulation of phenolic compounds in fresh produce in response to postharvest treatment with ET or MeJA has been previously reported (Heredia and Cisneros-Zevallos, 2009a). In the specific case of broccoli, the pre-harvest application of MeJA has been shown to increase the concentration of total glucosinolates in broccoli florets (Ku et al., 2013a,b). Additionally, previous reports have shown that MeJA may induce the accumulation of total phenolics and glucosinolates in broccoli sprouts (Pérez-Balibrea et al., 2011). However, the effect of postharvest treatments with exogenous ET and MeJA on the accumulation of glucosinolates and phenolic compounds in broccoli has not been reported.

Recently, it was reported that wounding triggers the production of reactive oxygen species (ROS), which induce the activation of primary and secondary metabolism in plants (Jacobo-Velázquez et al., 2015). Furthermore, the authors reported that other signaling molecules, such as ET and jasmonic acid (JA), which are produced after wounding, play key roles as ROS levels modulators, and mediate the expression of secondary metabolism genes, triggering the accumulation of specific secondary metabolites. Therefore, the accumulation of antioxidants in plants has also been studied as an effect of wounding during postharvest (Reyes and Cisneros-Zevallos, 2003; Jacobo-Velázquez et al., 2011; Surjadinata and Cisneros-Zevallos, 2012; Torres-Contreras et al., 2014a,b). In the case of broccoli, it has been reported that wounding triggers the biosynthesis and accumulation of indolic glucosinolates (Verkerk et al., 2001). Abiotic stress has been previously evaluated as a strategy to enhance the nutritional content of broccoli samples. However, the study of postharvest treatments has been scarce and the effect of combined application of two or more postharvest abiotic stresses has not been thoroughly evaluated although previous reports have shown that the *de novo* biosynthesis of glucosinolates is more likely to occur during postharvest storage (Ku et al., 2013a).

Wounding has been proven to be one of the most effective postharvest abiotic stresses for the activation of the phenylpropanoid metabolic pathway in plants (Reyes and Cisneros-Zevallos, 2003; Surjadinata and Cisneros-Zevallos, 2012; Torres-Contreras et al., 2014a). In addition, the application of ET and MeJA in wounded broccoli samples and their effect on the accumulation of phenolic compounds and glucosinolates has not been evaluated yet. Therefore, the objective of this research was to evaluate the effect of wounding stress alone and in combination with exogenous ET or MeJA, on the accumulation of total and individual phenolic compounds and glucosinolates during storage (24 h at 20°C) of broccoli tissue. Although it is well known that ET induces chlorophylls degradation in broccoli (Tian et al., 1994), in the present study quality parameters were not considered relevant to evaluate, since the objective was to find alternative uses to broccoli not intended for human consumption, such as biofactory of high value phenolic and

glucosinolate compounds. The stressed-broccoli tissue could be subjected to downstream processing in order to extract and purify bioactive molecules with applications in the dietary supplements, agrochemical and cosmetics markets.

MATERIALS AND METHODS

Chemicals

Sulfatase from *Helix pomatia*, sinigrin hydrate, sephadex A-25, 3-*O*-caffeoylquinic acid (3-*O*-CQA), acetonitrile (HPLC grade), methanol (HPLC grade), sodium acetate, MeJA, and orthophosphoric acid were obtained from Sigma Chemical Co. (St. Louis, MO, USA). ET was purchased from Infra (Naucalpan, MEX, Mexico). Acetic acid was purchased from Desarrollo de Especialidades Quimicas (Monterrey, NL, Mexico). Desulfoglucoraphanin was obtained from Santa Cruz Biotechnology (Dallas, TX, USA).

Plant Material, Processing, and Storage Studies

Broccoli (*Brassica oleracea*) was obtained on June 2014, once from a local market (HEB, Monterrey, Mexico), washed, and disinfected with chlorinated water (200 ppm, pH 6.5). All samples were supplied by the same grower. Whole heads were used as control samples for wounding stress. Florets and wounded florets (cut into four even pieces with a commercial straight-edged knife) were used as wounding treatments. Wounded and whole samples were stored inside hermetic plastic containers with periodic ventilation to avoid CO_2 accumulation over 0.5% (v/v). One broccoli head was used per replica of each treatment. Three biological replicates were performed for each treatment.

Ethylene and MeJA were applied as reported by Heredia and Cisneros-Zevallos (2009a), where ET was directly injected into the plastic containers (1 mL/L) and MeJA was applied by wetting a Whatman No. 4 filter paper (Whatman Inc., Piscataway, NJ, USA) over a Petri dish (0.25 mL/L). All samples were stored for 24 h in an incubator (VWR, Radnor, PA, USA) at 20°C to determine the treatment that yields a maximum accumulation of phenolic compounds and glucosinolates. Collected samples were freeze-dried (Labconco, Kansas City, MO, USA) prior to extraction of phytochemicals. The concentration of individual phenolics and glucosinolates was determined before and after storage.

Analysis of Phenolic Compounds by High-Performance Liquid Chromatography-Diode Array Detection (HPLC-DAD) and HPLC-Electrospray Ionization (ESI)-MS^n

Extraction of Phenolic Compounds

For the chromatographic detection and quantification of individual phenolic compounds, a methanol extract was prepared. Broccoli powder (0.5 g) was homogenized with methanol (20 mL) using a tissuemizer (Advanced homogenizing system, VWR). Subsequently, homogenates were centrifuged (9000 x *g*, 1 h, 4°C). The clear supernatant was filtered using nylon membranes (0.45 μm, VWR) prior to injection to the chromatographic systems.

Analysis of Phenolic Compounds by HPLC-DAD and HPLC-ESI-MS^n

The identification and quantification of individual phenolic compounds were performed as described by Torres-Contreras et al. (2014a). Briefly, 10 μL of the extract were injected in the HPLC-DAD system (1260 Infinity, Agilent Technologies, Santa Clara, CA, USA). Separation was done on a 4.6 mm × 250 mm, 5 μm particle size, C18 reverse phase column (Luna, Phenomenex, Torrance, CA, USA). Mobile phases consisted of water (phase A) and methanol/water (60:40, v/v, phase B) both adjusted at pH 2.4 with orthophosphoric acid. The gradient solvent system was 0/100, 3/70, 8/50, 35/30, 40/20, 45/0, 50/0, and 60/100 (min/% phase A) at a constant flow rate of 0.8 mL/min. Phenolic compounds were detected at 320 nm. Chromatographic data of analyses was processed with OpenLAB CDS ChemStation software (Agilent Technologies, Santa Clara, CA, USA).

Mass spectra were obtained on a MS Finnigan LCQ Deca XP Max, Ion trap mass spectrophotometer coupled at the exit of the DAD and equipped with a *Z*-spray ESI source, and run by Xcalibur version 1.3 software (Thermo Finnigan, San Jose, CA, USA). Separations were conducted in a Phenomenex Synergi^{TM} 4 μm Hydro-RP 80 Å (2 mm × 150 mm) with a C18 guard column, and a flow rate of 200 L/min from the DAD eluent was directed to the ESI interface using a flow-splitter. Mobile phases were adjusted to pH 2.4 with formic acid. Nitrogen was used as desolvation gas at 275°C and a flow rate of 60 L/h, and helium was used as damping gas. ESI was performed in the negative ion mode using the following conditions: sheath gas (N_2), 60 arbitrary units; spray voltage, 1.5 kV; capillary temperature, 285°C; capillary voltage, 45.7 V; tube lens offset, 30 V.

Individual phenolics were identified on the basis of retention time, UV spectra, and their mass-to-charge ratio as compared with authentic standards and a previous report (Vallejo et al., 2003). For the quantification of phenolic compounds, a standard curve of 3-*O*-CQA was prepared in the range of 0.5–100 μM. The concentration of phenolics was expressed as mg of 3-*O*-CQA equivalents per kg of broccoli dry weight (DW).

Analysis of Glucosinolates by High-Performance Liquid Chromatography-Diode Array Detection (HPLC-DAD) and HPLC-Electrospray Ionization (ESI)-MS^n

Extraction and Desulfation of Glucosinolates

For the chromatographic determination of glucosinolates, extraction, and desulfation was done as described by Kiddle et al. (2001) with modifications described by Saha et al. (2012). Briefly, 10 mL of methanol:water (70:30, v:v), previously heated for 10 min at 70°C, were added to broccoli powder (0.2 g) followed by 50 μL of a 3 mM solution of sinigrin as internal standard.

Samples were vortexed and incubated at 70°C for 30 min to ensure myrosinase inactivation. The extracts were removed from the water bath, left to cool at room temperature and centrifuged (3000 × *g*, 5 min, 4°C).

Afterward, glucosinolates were desulfated and purified using disposable polypropylene columns (Thermo Fisher Scientific, Waltham, MA, USA). To prepare the columns, 0.5 mL of water were added, followed by 0.5 mL of prepared Sephadex A-25 and an additional 0.5 mL of water. Clear supernatant (3 mL) was added into a prepared column and was allowed to drip through slowly. Columns were washed with 2 × 0.5 mL of water followed by 2 × 0.5 mL of 0.02 M sodium acetate. Purified sulfatase (75 μL) was added to each sample and was left at room temperature overnight (12 h). Desulfoglucosinolates were eluted with a total of 1.25 mL of water (0.5 mL + 0.5 mL + 0.25 mL).

Analysis of Desulfoglucosinolates by HPLC-DAD and HPLC-ESI-MS^n

Determination of desulfoglucosinolates was performed as reported by Vallejo et al. (2003) with slight modifications. Chromatographic separations were done on the same chromatographic systems and reverse phase columns used for the analysis of phenolic compounds. Separation of desulfoglucosinolates in the HPLC-DAD system was achieved using water (phase A) and acetonitrile (phase B) as mobile phases with a flow rate of 1.5 mL/min and a gradient of 0/100, 28/80, 30/100 (min/% phase A) with an injection volume of 20 μL. All compounds were detected at 227 nm.

For the HPLC-ESI-MS^n analyses, the gradient of the solvent system used to obtain mass spectra was 0/99, 16/80, 18/10 (min/% phase A) and a flow rate of 350 μL/min. Nitrogen was used as desolvation gas at 275°C and a flow rate of 60 L/h, and helium was used as damping gas. ESI was performed in the negative ion mode using the following conditions: sheath gas (N_2), 60 arbitrary units; spray voltage, 5 kV; capillary temperature, 285°C; capillary voltage, 48.5 V; tube lens offset, 30 V.

Individual glucosinolates were identified on the basis of retention time, UV spectra, and their mass-to-charge ratio as compared with authentic standards and previous reports (Hansen et al., 1996; Vallejo et al., 2003; Barbieri et al., 2008; Miglio et al., 2008). A standard curve of desulfoglucoraphanin was prepared in the range of 0–700 μM for the quantification of glucosinolates. The concentration of individual glucosinolates was expressed as mmol of desulfoglucoraphanin equivalents per kg of broccoli.

Statistical Analysis

Replication was achieved by repeating treatment under the same conditions. All treatments were run concurrently. All reported data were pooled from repeated independent treatment. There were three replicates per treatment ($n = 3$). Statistical analyses were performed using the three replicates. Data represent the mean values of samples and their standard error. Analyses of variance (ANOVA) were conducted using JMP software version 9.0 (SAS Institute Inc., Cary, NC, USA) and mean separations performed using the LSD test ($p < 0.05$).

RESULTS AND DISCUSSION

Effect of Wounding Stress, Phytohormone Treatment, and Storage Time on the Accumulation of Phenolic Compounds

The identification of individual phenolic compounds present in broccoli treated with or without wounding and phytohormones is shown in **Figure 1** and **Table 1**. The chemical structure of individual phenolic compounds identified are shown in **Figure 2** (compounds **1**–**7**) and included 3-*O*-CQA (compound **1**), 5-*O*-caffeoylquinic acid (5-*O*-CQA, compound **2**), 1,2-disinapoylgentiobiose (1,2-DSG, compound **3**), 1-sinapoyl-2-ferulolylgentiobiose (1-S-2-FG, compound **4**), 1,2,2-trisinapoylgentiobiose (1,2,2-TSG, compound **5**),

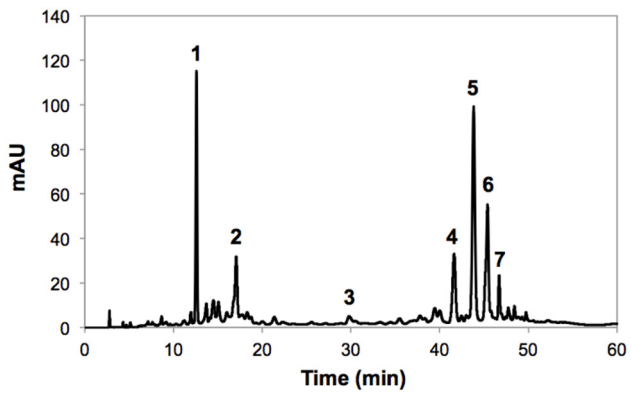


FIGURE 1 | Typical HPLC-DAD chromatogram (shown at 320 nm) from methanol extracts of broccoli whole heads before storage. Tentative identification of peaks was performed as indicated in **Table 1**. Peak assignment: (1) 3-*O*-CQA; (2) 5-*O*-CQA; (3) 1,2-DSG; (4) 1-S-2-FG (5) 1,2,2-TSG; (6) 1,2-DFG; (7) 1,2-DS-2-FG.

TABLE 1 | Tentative identification of individual phenolic compounds in broccoli.

Peak number (retention time, min)	λmax (nm)	Tentative identification	[M-H]$^-$ (m/z)
1 (12.8)	295, 332	3-*O*-CQA[a,b]	353
2 (17.2)	295, 332	5-*O*-CQA[b]	353
3 (30.1)	328	1,2-DSG[b]	753
4 (44.3)	295, 328	1-S-2-FG[b]	723
5 (45.8)	328	1,2,2-TSG[b]	959
6 (47.1)	290,328	1,2-DFG[b]	693
7 (48.2)	295, 328	1,2-DS-2-FG[b]	929

Identification was obtained by HPLC-DAD-ESI-MS^n.
[a]*Identified based on their spectra characteristics and their mass-to-charge ratio as compared with authentic standards.* [b]*Identified based on their spectra characteristics and their mass-to-charge ratio as compared with a previous report (Vallejo et al., 2003). 3-O-caffeoylquinic acid (3-O-CQA); 5-O-caffeoylquinic acid (5-O-CQA); 1,2-disinapoylgentiobiose (1,2-DSG); 1-sinapoyl-2-ferulolylgentiobiose (1-S-2-FG); 1,2,2-trisinapoylgentiobiose (1,2,2-TSG); 1,2-diferulolylgentiobiose (1,2-DFG); 1,2-disinapoyl-2-ferulolylgentiobiose (1,2-DS-2-FG).*

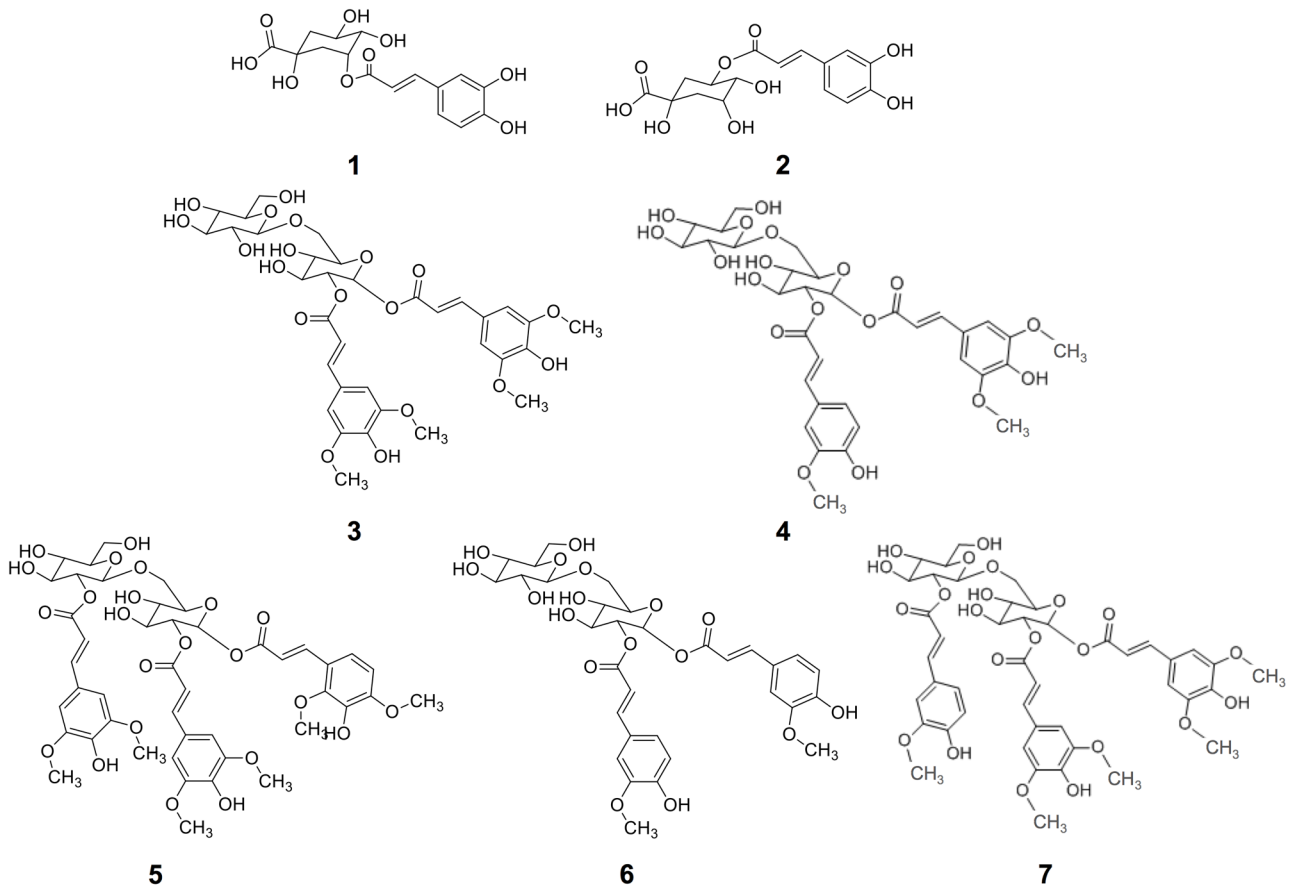


FIGURE 2 | Chemical structures of phenolic compounds identified in broccoli subjected to wounding stress and exogenous ET and MeJA: 5-*O*-caffeoylquinic acid (3-*O*-CQA, 1), 5-*O*-caffeoylquinic acid (5-*O*-CQA, 2), 1,2-disinapoylgentiobiose (1,2-DSG, 3), 1-sinapoyl-2-ferulolylgentiobiose (1-S-2-FG, 4), 1,2,2-trisinapoylgentiobiose (1,2,2-TSG, 5), 1,2-diferulolylgentiobiose (1,2-DFG, 6), and 1,2-disinapoyl-2-ferulolylgentiobiose (1,2-DS-2-FG, 7).

1,2-diferulolylgentiobiose (1,2-DFG, compound **6**), and 1,2-disinapoyl-2-ferulolylgentiobiose (1,2-DS-2-FG, compound **7**). Phenolic compounds identified in broccoli samples herein evaluated (**Figures 1** and **2**; **Table 1**) agree with a previous report (Vallejo et al., 2003). However, in the present study 1,2,2-TSG was identified as the major phenolic compound in broccoli instead of 3-*O*-CQA. The observed differences in phenolic profiles may be due to different cultivation conditions and genetic variation, although extraction parameters are also likely to influence the quantification of phytochemicals (Luthria, 2012).

The phenolic profile of broccoli was affected by the treatments applied (**Table 2**). In the specific case of broccoli heads, storage at 20°C for 24 h resulted on higher levels of phenolics, which were reflected on ∼33, ∼30, and 46% higher levels of 1,2,2-TSG, 1,2-DFG, and 1,2-DS-2-FG, respectively, while the content of 3-*O*-CQA, 5-*O*-CQA, 1,2-DSG, and 1-S-2-FG was not affected by storage. Similar observations have been previously reported in the literature. For instance, Starzyńska et al. (2003) reported ∼22% higher levels of total phenols in broccoli heads stored for 24 h at 20°C. Similarly, Costa et al. (2006) showed that broccoli heads stored for 2 days at 20°C had ∼36% more total phenols than samples before storage. These increments in phenolics observed after storage of broccoli heads could be attributed to a transient increase in the activity of phenylalanine ammonia-lyase (PAL), a key enzyme involved in the biosynthesis of phenolic compounds, as has been previously reported during the first 12 h of storage of broccoli heads stored at 20°C (Porras-Baclayon et al., 2007).

On the other hand, when exogenous ET was applied to whole heads, the accumulation of 1,2,2-TSG was inhibited, whereas MeJA induced a decrease by ∼19% in the content of 3-*O*-CQA as compared to time 0 h control samples, and repressed the accumulation of 1-S-2-FG, 1,2,2-TSG, 1,2-DFG, and 1,2-DS-2-FG (**Table 2**). Although the effects of postharvest treatments with ET and MeJA on the accumulation of phenols in broccoli heads has not been previously studied, Heredia and Cisneros-Zevallos (2009a) reported that the postharvest exposure of whole fresh produce, such as asparagus, potatoes, apples, peaches, strawberries, and grapes, to ET and MeJA had no effect on the total phenolic content of each crop after four days of storage at 20°C. Additionally, other reports by the same authors showed that the concentration of total phenolics in whole carrots treated with ET or MeJA remained

TABLE 2 | Concentration of individual phenolic compounds in whole and wounded broccoli tissue treated with and without ethylene (ET) and methyl jasmonate (MeJA) and stored for 24 h at 20°C.

Treatment	Sample	Individual phenolic concentration (mg/kg)[1,2,3]						
		3-*O*-CQA	5-*O*-CQA	1,2-DSG	1-S-2-FG	1,2,2-TSG	1,2-DFG	1,2-DS-2-FG
Control (0 h)	Whole	266.57 ± 8.56 bc	124.07 ± 2.99 a	20.19 ± 2.37 c	164.04 ± 6.97 ab	432.18 ± 20.75 bcde	229.99 ± 20.75 bcd	45.36 ± 3.28 ef
Air Control(24 h)	Whole	270.40 ± 5.65 bc	121.19 ± 2.86 a	29.24 ± 2.50 c	199.62 ± 7.48 a	575.40 ± 41.76 a	300.26 ± 43.75 a	66.54 ± 9.35 abc
	Floret	326.75 ± 6.04 a	109.35 ± 2.58 a	57.65 ± 3.14 a	164.85 ± 50.73 ab	487.10 ± 6.10 b	200.01 ± 11.74 cde	74.89 ± 4.82 a
	Wounded floret	242.20 ± 5.70 cd	106.46 ± 13.58 a	24.32 ± 8.25 c	49.68 ± 36.35 d	395.62 ± 26.75 de	220.40 ± 11.35 bcde	53.25 ± 1.10 de
Ethylene(24 h)	Whole	281.52 ± 11.87 b	124.75 ± 8.84 a	16.44 ± 3.21 c	171.31 ± 12.47 ab	475.72 ± 31.23 bc	237.45 ± 37.86 abcd	61.20 ± 3.63 bcd
	Floret	287.37 ± 7.53 b	78.21 ± 6.71 b	50.56 ± 3.34 b	157.22 ± 5.60 bc	492.02 ± 23.86 b	185.01 ± 12.42 cde	62.81 ± 4.36 bcd
	Wounded floret	252.35 ± 19.80 bcd	119.94 ± 9.18 a	37.11 ± 8.08 c	151.08 ± 6.22 bc	445.21 ± 25.59 bcd	273.24 ± 31.16 ab	69.67 ± 9.22 ab
Methyl Jasmonate (24 h)	Whole	215.48 ± 8.60 d	119.90 ± 1.49 a	14.32 ± 1.02 c	154.65 ± 9.43 bc	430.58 ± 14.63 bcde	169.71 ± 9.80 de	38.33 ± 7.72 f
	Floret	282.52 ± 15.45 b	56.18 ± 2.71 c	50.64 ± 3.19 b	123.85 ± 8.86 c	375.17 ± 22.61 de	152.17 ± 5.49 e	57.19 ± 4.35 cd
	Wounded floret	242.81 ± 24.64 cd	114.27 ± 3.86 a	18.46 ± 0.20 c	139.54 ± 3.59 bc	409.66 ± 21.39 cde	240.13 ± 20.55 abc	60.86 ± 10.24 bcd

[1]*Concentrations are reported as 3-O-CQA equivalents. All compounds were quantified at 320 nm.* [2]*Values represent the mean of three replicates ± standard error of the mean.* [3]*Different letters in the same column indicate statistical difference by the LSD test ($p < 0.05$). 3-O-caffeoylquinic acid (3-O-CQA); 5-O-caffeoylquinic acid (5-O-CQA); 1,2-disinapoylgentiobiose (1,2-DSG); 1-sinapoyl-2-feruloylgentiobiose (1-S-2-FG); 1,2,2-trisinapoylgentiobiose (1,2,2-TSG); 1,2-diferulolylgentiobiose (1,2-DFG); 1,2-disinapoyl-2-feruloylgentiobiose (1,2-DS-2-FG).*

unchanged throughout 12 days of storage at 15°C (Heredia and Cisneros-Zevallos, 2009b). However, in both cases, whole tissues stored under air control conditions did not show an enhanced content of total phenolic compounds. This particular difference between the results obtained herein and previous reports may be due to a variation among crops in the response to the stress caused by storage conditions (time and temperature). Moreover, Yuan et al. (2010) reported that the treatment of broccoli florets with 1-methylcyclopropene (1-MCP), an inhibitor of ET action, induced an increase on the activity of the enzyme superoxide dismutase, which produces hydrogen peroxide (H_2O_2), a signaling molecule involved on the activation of PAL gene expression and enzymatic activity (Jacobo-Velázquez et al., 2011). Besides, Jacobo-Velázquez et al. (2015) reported a repression in ROS levels as a response to ET treatment in shredded carrots. Therefore, it is likely that broccoli heads treated with exogenous ET could present lower levels of ROS, leading to lower activation of PAL and lower accumulation of total phenolics.

Regarding the concentration of phenolic compounds in broccoli florets before and after storage under air control conditions, the content of 3-*O*-CQA, 1,2-DSG, and 1,2-DS-2-FG showed an increase of ∼22, ∼185, and ∼65%, respectively, as compared to time 0 h control samples, whereas the concentration of 5-*O*-CQA, 1,2-DFG, 1-S-2-FG, and 1,2,2-TSG remained unaltered (**Table 2**). In agreement with this observation, a previous report showed that the total phenolic content remained unchanged during the first twelve days of storage of broccoli florets stored at 5°C (Amodio et al., 2014). Likewise, Du et al. (2014) reported that broccoli florets stored for three days at 15°C showed a slight but significant increase in total soluble phenols (1.1-fold) due to wounding. Chlorogenic acid (3-*O*-CQA) is one of the principal precursors of lignin, whereas 1,2-DSG, and 1,2-DS-2-FG are glycosides of sinapic and ferulic acid, which aglycones are also utilized for the biosynthesis of coniferyl alcohols, precursors of lignin (Vanholme et al., 2010; Torres-Contreras et al., 2014a,b). Therefore, the higher levels of 3-*O*-CQA, 1,2-DSG, and 1,2-DS-2-FG observed after storage of broccoli florets under air control conditions may be related with the wound-induced activation of the phenylpropanoid metabolism, which is required for the biosynthesis of lignin that in wounded plant tissue serves as a water impermeable barrier that prevent excessive water loss (Whetten and Sederoff, 1995; Boerjan et al., 2003; Becerra-Moreno et al., 2015).

The application of ET to broccoli florets during storage inhibited the accumulation of 3-*O*-CQA, and 1,2-DS-2-FG, whereas the accumulation of 1,2-DSG was decreased by ∼12% as compared to air control florets (**Table 2**). Likewise, ET treated broccoli florets showed ∼36% lower levels of 5-*O*-CQA as compared with samples before storage (**Table 2**). It has been reported that 1-MCP induces the downregulation in the expression of genes related with lignin biosynthesis in *Brassica chinensis*, while ET upregulates them (Zhang et al., 2010). Therefore, the less intense accumulation of 1,2-DSG and 1,2-DS-2-FG, as well as the impeded accumulation of 3-*O*-CQA and lower levels of 5-*O*-CQA observed in ET treated florets, suggests an increased rate of lignin biosynthesis induced by ET. Florets

treated with MeJA showed a more moderate accumulation of 1,2-DSG and 1,2-DS-2-FG and a decrease of 5-*O*-CQA as compared with the air control (**Table 2**). Additionally, lower concentrations of 1-S-2-FG and 1,2-DFG were observed. Previous reports showed that JA downregulates genes involved in the biosynthesis of phenolic compounds, such as *PAL* and *4-coumarate-CoA ligase* (*4CL*), as well as genes involved in the biosynthesis of lignin, such as the *caffeoyl-CoA 3-O-methyltransferase* (*CCoAOMT*) gene (Jacobo-Velázquez et al., 2015). Therefore, unlike ET-treated samples, the observed MeJA induced decrease and repression of accumulation of individual phenolic compounds are likely due to a downregulation of genes involved in the secondary metabolic pathways leading to the biosynthesis of phenolic compounds.

The application of additional wounding stress to broccoli florets (florets cut into four even pieces) induced a decrease in concentration of 1-S-2-FG by ∼69% after 24 h of storage, whereas the concentration of the other phenolics remained unchanged (**Table 2**). As earlier described, the decrement in phenolics induced by wounding could be attributed to their conversion into lignin, which is needed to prevent excessive water loss in wounded plants (Whetten and Sederoff, 1995; Boerjan et al., 2003; Becerra-Moreno et al., 2015). Therefore, it is likely that wounded broccoli florets experienced a higher rate of lignification than the rate of phenolics biosythesis, and thus, lower levels of 1-S-2-FG were detected. The phenolic compounds identified in broccoli have a glycosylated structure in which two or three simple phenolics may be attached (compounds **3–7**, **Figure 2**). In the case of wounded florets stored under air control conditions, the main phenolic compound affected by wounding (1-S-2-FG) has one sinapic acid and one ferulic acid attached to the carbohydrate moiety. Therefore, it could be hypothesized that phenolic glycosides with a lower number of simple phenolics in their structures are more prone to be hydrolyzed and used as lignin building blocks.

ET applied to wounded-florets resulted on ∼53% higher levels of 1,2-DS-2-FG as compared to samples before storage and ∼30% higher content than air control wounded florets. Furthermore, the application of ET in wounded florets impeded the decrease in concentration observed in 1-S-2-FG after 24 h of storage of wounded florets (**Table 2**). As described earlier, previous reports indicates that ET activates the expression of genes related with phenolics and lignin biosynthesis in wounded plants (Jacobo-Velázquez et al., 2015). In the specific case of wounded carrots, the application of exogenous ET to the tissue increased PAL activity and phenolics accumulation (Heredia and Cisneros-Zevallos, 2009b). Therefore, it is likely that in wounded broccoli florets, exogenous ET increased the biosynthesis rate of phenolic compounds as compared to the air control, and since phenolic biosynthesis increased, the balance between phenolic production and utilization for lignin biosynthesis resulted in no change in total phenolic content.

The application of MeJA to wounded florets only affected the concentration of 1,2-DS-2-FG as compared to air control samples, where exogenous MeJA impeded the wound-induced decrease in concentration observed after storage. These results are in agreement with a previous report, where the application of MeJA to wounded carrots stored for 12 day at 15°C did not induce a significant increase in the concentration of total phenolics (Heredia and Cisneros-Zevallos, 2009b). Additionally, pre-harvest studies have shown that the concentration of phenolic compounds in broccoli is not affected in response to treatment with MeJA. For instance, Barrientos-Carvacho et al. (2014) reported that the treatment of broccoli sprouts with three different concentrations of MeJA (10, 50, 90 μM) induced a decrease in total phenolic compounds. Similarly, a study by Ku and Juvik (2013) showed that the application of MeJA (250 μM) to aerial tissues of broccoli 4 days prior to harvest had no effect on the concentration of phenolic compounds in broccoli florets. These observations may be due to the downregulation that JA exerts on genes related with phenolics and lignin biosynthesis in wounded plants (Jacobo-Velázquez et al., 2015). Therefore, results presented herein indicate that MeJA is not an elicitor for the phenylpropanoid pathway in broccoli, however, since 1-S-2-FG content was increased in wounded-florets treated with MeJA as compared with the air control, it is likely that MeJA selectively induce the accumulation of 1-S-2-FG in wounded tissue (**Table 2**).

Given their health-promoting properties, the production of phenolic compounds in broccoli would be of great interest for the pharmaceutical and dietary supplements industry. For instance, 3-*O*-caffeoylquinic acid has been associated with the reduction of the risk of developing cardiovascular diseases, type II diabetes, and neurodegenerative diseases (Farah et al., 2008). Furthermore, ferulic acid and sinapic acid, the phenolic aglycones of 1,2-DSG, 1-S-2-FG, 1,2,2-TSG, 1,2-DFG, and 1,2-DS-2-FG, are important antioxidants that inhibit the peroxidation of LDL, helping to prevent the progression of atherosclerosis (Natella et al., 1999).

Effect of Wounding Stress, Phytohormone Treatment, and Storage Time on the Accumulation of Glucosinolates

The identification of individual glucosinolates present in broccoli treated with or without wounding and phytohormones is shown in **Figure 3** and **Table 3**. The chemical structure of individual glucosinolates identified is shown in **Figure 4** (compounds **1–4**) and includes one aliphatic glucosinolate (glucoraphanin, compound **1**), and three indolic glucosinolates (4-hydroxyglucobrassicin, compound **2**; glucobrassicin, compound **3**; and neoglucobrassicin, compound **4**). Likewise, their concentration in broccoli treated with or without wounding and phyhormones is shown in **Figure 5**.

Whole heads stored for 24 h at 20°C showed an increase of ∼84% in the content of 4-hydroxyglucobrassicin as compared to the control (time 0 h samples), whereas the content of the other three individual glucosinolates remained unchanged (**Figure 5**). When whole broccoli heads were treated with ET, the concentration of glucoraphanin and 4-hydroxyglucobrassicin increased by ∼52 and ∼223%, respectively, as compared with the control (**Figures 5A,C**). These results are in agreement with a previous report, where the expression levels of genes related with glucosinolate biosynthesis strongly correlated with endogenous ET production in broccoli (Ku et al.,

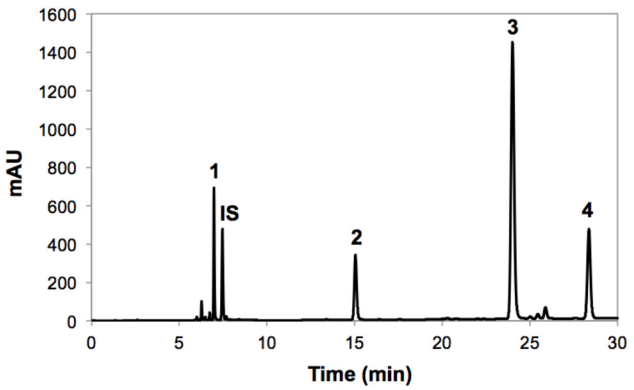


FIGURE 3 | Typical HPLC-DAD chromatogram (shown at 227 nm) from methanol/water (70/30, v/v) extracts of broccoli whole heads before storage. Tentative identification of peaks was performed as indicated in **Table 3**. Peak assignment: (1) Glucoraphanin; (2) 4-hydroxyglucobrassicin; (3) Glucobrassicin; (4) Neoglucobrassicin; (IS) Internal standard (sinigrin).

TABLE 3 | Identification of individual glucosinolates in broccoli.

Peak number (retention time, min)	λmax (nm)	Tentative identification	[M-H]$^-$ (m/z)
1 (7.1)	222	Glucoraphanin[a,b,c]	380
2 (14.8)	221	4-hydroxyglucobrassicin[b]	–
3 (23.1)	222, 280	Glucobrassicin[b,c]	391
4 (27.5)	222	Neoglucobrassicin[b,c]	421

Identification was obtained by HPLC-PDA-ESI-MS^n.
[a]*Identified based on their spectra characteristics and their mass-to-charge ratio as compared with authentic standards.* [b]*Identified based on their spectra characteristics and order of elution as compared with a previous report (Hansen et al., 1996; Vallejo et al., 2003; Barbieri et al., 2008; Miglio et al., 2008).* [c]*Identified based on their spectra characteristics and their mass-to-charge ratio as compared with a previous report (Vallejo et al., 2003).*

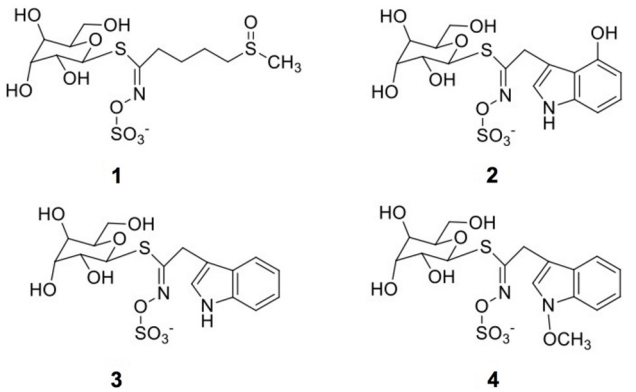


FIGURE 4 | Chemical structures of glucosinolates identified in broccoli subjected to wounding stress and exogenous ET and MeJA: glucoraphanin (1), 4-hydroxyglucobrassicin (2), glucobrassicin (3), neoglucobrassicin (4).

2013b). In the indolic glucosinolate biosynthetic pathway, glucobrassicin is synthesized by sulfotransferases 16 and 18 (SOT16 and SOT18), and then glucobrassicin is converted into neoglucobrassicin and 4-hydroxyglucobrassicin by the subfamily of CYP81F genes by methylation and hydroxylation reactions, respectively. Therefore, these observations suggest that, in the specific case of broccoli whole heads, the hydroxylation of glucobrassicin was favored by postharvest storage and ET treatments.

The application of MeJA in whole broccoli heads did not induce additional accumulation of glucosinolates as compared to the air control (**Figure 5**). A previous report where MeJA was applied four days before harvest of broccoli reported an increase in the expression levels of hydrolytic enzymes (myrosinase), which converts glucosinolates into isothiocyanates (Ku et al., 2013b). This suggests that MeJA also acts as a signal that leads to a higher myrosinase activity in whole heads, making it more available for glucosinolate hydrolysis, therefore, it is likely that the glucoraphanin and 4-hydroxyglucobrassicin produced by storage conditions are being hydrolyzed into isothiocyanates, and thus, no accumulation of any individual glucosinolate was observed (**Figure 5**).

Florets stored for 24 h at 20°C under air control conditions did not show significant difference in the glucosinolate profiles as compared with the control (time 0 h samples, **Figure 5**). Treating broccoli florets with ET or MeJA did not affect the concentration of individual glucosinolates as compared with air control samples. The application of additional wounding stress to florets (wounded florets) did not affect the glucosinolate profile of broccoli when stored under air control conditions. However, when wounded florets were treated with ET, the concentration of neoglucobrassicin and 4-hydroxyglucobrassicin was enhanced by ∼193 and ∼117% as compared to the control (time 0 h samples), whereas the concentration of the other individual glucosinolates remained unaltered as compared with the air control. Likewise, MeJA treatments induced significant increments in the levels of neoglucobrassicin and 4-hydroxyglucobrassicin, where their concentrations were increased by ∼286 and ∼117% as compared with time 0 h samples. Results suggest that MeJA and ET induced the activation of CYP81F genes involved on glucobrassicin hydroxylation and methoxylation, forming 4-hydroxyglucobrassicin and neoglucobrassicin, respectively. These results are in agreement with studies in *Arabidopsis* indicating that genes related to indolic glucosinolate biosynthesis are more susceptible to be induced by exogenous phytohormones rather than those playing a role in aliphatic glucosinolate biosynthesis, although the latter may also be up-regulated (Mikkelsen et al., 2003). Interestingly, in the present study, it was shown that to accumulate indolic glucosinolates (neoglucobrassicin and 4-hydroxyglucobrassicin) the combination of wounding stress with MeJA or ET was required, whereas for the accumulation of aliphatic glucosinolate (glucoraphanin) the sole application of ET on broccoli heads was sufficient (**Figure 5**). The unchanged levels of glucoraphanin in response to exogenous phytohormones in broccoli treated

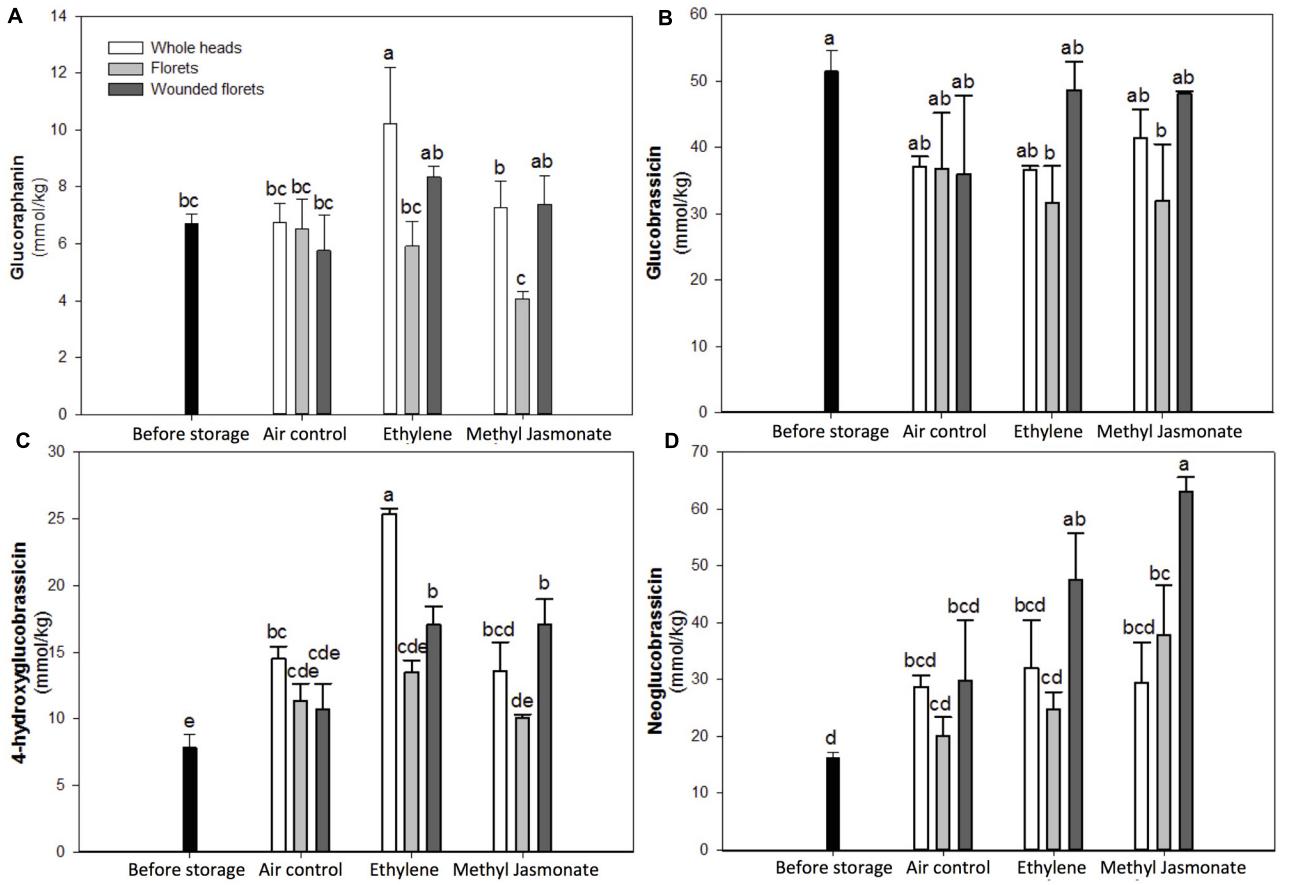


FIGURE 5 | Effect of wounding and exogenous ET and MeJA on the concentration of glucoraphanin (A), glucobrassicin (B), 4-hydroxyglucobrassicin (C), and neoglucobrassicin (D) of broccoli stored for 24 h at 20°C. Data represent the mean of three repetitions and their standard errors. Data reported for samples before storage represent values of time 0 h whole heads. Bars with different letters indicate statistical difference by the LSD test ($p < 0.05$).

with wounding stress (florets and wounded florets) may be due to a selective wound-induced activation of an aliphatic-specific myrosinase.

The accumulated glucosinolates in stressed broccoli have diverse industrial applications. For instance, in the pharmaceutical and dietary supplements industry, glucoraphanin has gained interest in the last few years due to the anticarcinogenic properties of its hydrolysis product, sulforaphane (Fahey and Talalay, 1999). Additionally, glucosinolates can also be used as insecticides to protect horticultural crops, as reported by El Sayed et al. (1996), who showed an inhibiting effect of glucobrassicin and its hydrolysis product, indol-3-ylmethylisothiocyanate, on *Schistocerca gregaria*, an insect that threatens crop production mainly in Africa, Middle East, and Asia.

CONCLUSION

Results presented herein showed that simple postharvest treatments such as wounding applied alone or in combination with exogenous phytohoromnes (ET and MeJA) can be used as an effective emerging technology that allows the accumulation of specific glucosinolate and phenolic compounds in broccoli. For instance if the accumulation of specific phenolic compounds such as 2,2-TSG, 1,2-DFG, and 1,2-DS-2- FG are desired whole broccoli heads can be stored for 24 h at 20°C. Furthermore, for the accumulation of 3-*O*-CQA, 1,2-DSG, and 1,2-DS-2-FG, broccoli florets should be stored under the same condition. However, when broccoli was treated with ET or MeJA the accumulation of these individual phenolics was impeded. On the other hand, if the accumulation of glucoraphanin and 4-hydroxyglucobrassicin is desirable, whole broccoli heads should be treated with exogenous ET for 24 h. Likewise, for the accumulation of neoglucobrassicin wounded broccoli florets should be treated with exogenous ET and MeJA during 24 h at 20°C. This particular observation suggests a complex cross-talk between wounding and the applied phytoregulators acting on the metabolism of glucosinolates. Despite that quality changes in broccoli were not a factor evaluated in this study, the visual quality of broccoli heads was not affected within the 24-h period of storage, suggesting that it could be used as a functional food. However, further studies should be performed to validate

consumers acceptability and microbial safety of the tissue. Additionally, the stressed broccoli tissue with increased levels of bioactive molecules could be subjected to downstream processing in order to extract and purify the bioactive compounds for their subsequent use on the dietary supplements, agrochemical and cosmetics markets.

AUTHOR CONTRIBUTIONS

DV-G, and DJ-V designed experiments. DV-G and VN, carried out experiments. DV-G and VN, processed data. DV-G, LC-Z, and DJ-V, analyzed data and wrote the main text of the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This study is based upon research supported by research funds from Consejo Nacional de Ciencia y Tecnología (CONACYT, México) Grant (177012), CONACYT and Texas A&M University (CONACYT-TAMU) agreement Project number 2014-032(S) and Tecnologico de Monterrey (Bioprocess and Synthetic biology Research Group). Author DV-G also acknowledges the scholarship (296572) from CONACYT.

REFERENCES

Agricultural Marketing Resource Center [AgMRC] (n.d.). Available at: http://www.agmrc.org/ [accessed April 13, 2015].

Amodio, M. L., Derossi, A., and Colelli, G. (2014). Modeling phenolic content during storage of cut fruit and vegetables: a consecutive reaction mechanism. *J. Food Eng.* 140, 1–8. doi: 10.1016/j.jfoodeng.2014.04.006

Barbieri, G., Pernice, R., Maggio, A., De Pascale, S., and Fogliano, V. (2008). Glucosinolates profile of *Brassica rapa* L. subsp. *Sylvestris* L. Janch. Var. esculenta Hort. *Food Chem.* 107, 1687–1691. doi: 10.1016/j.foodchem.2007.09.054

Barrientos-Carvacho, H., Pérez, C., Zúñiga, G., and Mahn, A. (2014). Effect of methyl jasmonate, sodium selenate, and chitosan as exogenous elicitors on the phenolic compounds profile of broccoli sprouts. *J. Sci. Food Agric.* 94, 2555–2561. doi: 10.1002/jsfa.6596

Becerra-Moreno, A., Redondo-Gil, M., Benavides, J., Nair, V., Cisneros-Zevallos, L., and Jacobo-Velázquez, D. (2015). Combined effect of water loss and wounding stress on gene activation of metabolic pathways associated with phenolic biosynthesis in carrot. *Front. Plant Sci.* 6:837. doi: 10.3389/fpls.2015.00837

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–546. doi: 10.1146/annurev.arplant.54.031902.134938

Brewer, M. S. (2011). Natural antioxidants: sources, compounds, mechanisms of action, and potential applications. *Compr. Rev. Food Sci. Food Safe.* 10, 221–247. doi: 10.1111/j.1541-4337.2011.00156.x

Cisneros-Zevallos, L. (2003). The use of controlled postharvest abiotic stresses as a tool for enhancing the nutraceutical content and adding-value of fresh fruits and vegetables. *J. Food Sci.* 68, 1560–1565. doi: 10.1111/j.1365-2621.2003.tb12291.x

Costa, L., Vicente, A. R., Civello, P. M., Chaves, A. R., and Martínez, G. A. (2006). UV-C treatment delays postharvest senescence in broccoli florets. *Postharvest Biol. Tecnol.* 39, 204–210. doi: 10.1016/j.postharvbio.2005.10.012

Del Rio, D., Rodriguez-Mateos, A., Spencer, J. P. E., Tognolini, M., Borges, G., and Crozier, A. (2013). Dietary polyphenolics in human health: structures, bioavailability, and evidence of protective effects against chronic diseases. *Antioxd. Redox Signal.* 18, 1818–1892. doi: 10.1089/ars.2012.4581

Du, W.-X., Avena-Bustillos, R. J., Breska, A. P. III, and McHugh, T. H. (2014). UV-B light as a factor affecting total soluble phenolic contents of various whole and fresh-cut specialty crops. *Postharvest Biol. Tecnol.* 93, 72–82. doi: 10.1016/j.postharvbio.2014.02.004

Economics, Statistics, and Market Information System [ERS] (n.d.). Available at: http://www.ers.usda.gov [accessed April 27, 2015].

El Sayed, G., Louveaux, A., Mavratzotis, M., Rollin, P., and Quinsac, A. (1996). Effects of glucobrassicin, epiprogoitrin and related breakdown products on locusts feeding: schouwia purpurea and desert locust relationships. *Entomol. Exp. Appl.* 78, 231–236. doi: 10.1111/j.1570-7458.1996.tb00786.x

Fahey, J. W., Haristoy, X., Dolan, P. M., Kensler, T. W., Scholtus, I., Stephenson, J. J., et al. (2002). Sulforaphane inhibits extracellular, intracellular, and antibiotic-resistant strains of *Helicobacter* pylori and prevents benzo[a]pyrene-induced stomach tumors. *Proc. Natl. Acad. Sci. U.S.A.* 99, 7610–7615. doi: 10.1073/pnas.112203099

Fahey, J. W., and Talalay, P. (1999). Antioxidant functions of sulforaphane: a potent inducer of Phase 2 detoxication enzymes. *Food Chem. Toxicol.* 37, 973–979. doi: 10.1016/S0278-6915(99)00082-4

Farah, A., Monteiro, M., Donangelo, C. M., and Lafay, S. (2008). Chlorogenic acids from green coffee extract are highly bioavailable in humans. *J. Nutr.* 138, 2309–2315. doi: 10.3945/jn.108.095554

Food and Agricultural Organization of the United Nations [FAOSTAT] (n.d.). Available at: http://faostat.fao.org/ [accessed April 27, 2015].

Food and Agriculture Organization of the United Nations [FAO] (n.d.). Available at: http://www.fao.org [accessed May 14, 2015].

Hansen, M., Møller, P., Sørensen, H., and Cantwell de Trejo, M. (1996). Glucosinolates in broccoli stored under controlled atmosphere. *J. Am. Soc. Hortic. Sci.* 120, 1069–1074.

Heo, H. J., Kim, Y. J., Chung, D., and Kim, D. O. (2007). Antioxidant capacities of individual and combined phenolics in a model system. *Food Chem.* 104, 87–92. doi: 10.1016/j.foodchem.2006.11.002

Heredia, J. B., and Cisneros-Zevallos, L. (2009a). The effects of exogenous ethylene and methyl jasmonate on the accumulation of phenolic antioxidants in selected whole and wounded fresh produce. *Food Chem.* 115, 1500–1508. doi: 10.1016/j.foodchem.2009.01.078

Heredia, J. B., and Cisneros-Zevallos, L. (2009b). The effect of exogenous ethylene and methyl jasmonate on pal activity, phenolic profiles and antioxidant capacity of carrots (Dacus carota) under different wounding intensities. *Postharvest Biol. Chem.* 51, 242–249. doi: 10.1016/j.postharvbio.2008.07.001

Jacobo-Velázquez, D. A., and Cisneros-Zevallos, L. (2009). Correlations of antioxidant activity against phenolic content revisited: a new approach in data analysis for food and medicinal plants. *J. Food Sci.* 74, 107–113. doi: 10.1111/j.1750-3841.2009.01352.x

Jacobo-Velázquez, D. A., and Cisneros-Zevallos, L. (2012). An alternative use of horticultural crops: stressed plants as biofactories of bioactive phenolic compounds. *Agriculture* 2, 259–271. doi: 10.3390/agriculture2030259

Jacobo-Velázquez, D. A., González-Agüero, M., and Cisneros-Zevallos, L. (2015). Cross-talk between signaling pathways: the link between plant secondary metabolite production and wounding stress response. *Sci. Rep.* 5:8608. doi: 10.1038/srep08608

Jacobo-Velázquez, D. A., Martínez-Hernández, G. B., Rodríguez, S. C., Cao, C. M., and Cisneros-Zevallos, L. (2011). Plants as biofactories: physiological role of reactive oxygen species on the accumulation of phenolic antioxidants in carrot tissue under wounding and hyperoxia stress. *J. Agric. Food Chem.* 59, 6583–6593. doi: 10.1021/jf2006529

Juurlink, B. (2001). Therapeutic potential of dietary phase 2 enzyme inducers in ameliorating diseases that have an underlying inflammatory component. *Can. J. Physiol. Pharm.* 79, 266–282. doi: 10.1139/y00-120

Kiddle, G., Bennett, R. N., Botting, N. P., Davidson, N. E., Robertson, A. A. B., and Wallsgrove, R. M. (2001). High-performance liquid chromatographic separation of natural and synthetic desulphoglucosinolates and their chemical

validation by UV, NMR and chemical ionisation-MS methods. *Phytochem. Anal.* 12, 226–242. doi: 10.1002/pca.589

Kim, H. S., and Juvik, J. A. (2011). Effect of selenium fertilization and methyl jasmonate treatment on glucosinolate accumulation in broccoli florets. *J. Amer. Soc. Hort. Sci.* 136, 239–246.

Ku, K. M., Choi, J. H., Kim, H. S., Kushad, M. M., Jeffery, E. H., and Juvik, J. A. (2013a). Methyl jasmonate and 1-methylcyclopropene treatment effects on quinone reductase inducing activity and post-harvest quality of broccoli. *PLoS ONE* 8:e77127. doi: 10.1371/journal.pone.0077127

Ku, K. M., Jeffery, E. H., and Juvik, J. A. (2013b). Influence of seasonal variation and methyl jasmonate mediated induction of glucosinolate biosynthesis on quinone reductase activity in broccoli florets. *J. Agric. Food Chem.* 61, 9623–9631. doi: 10.1021/jf4027734

Ku, K. M., Jeffery, E. H., and Juvik, J. A. (2014). Optimization of methyl jasmonate application to broccoli florets to enhance health-promoting phytochemical content. *J. Sci. Food Agric.* 94, 2090–2096. doi: 10.1002/jsfa.6529

Ku, K. M., and Juvik, J. A. (2013). Environmental stress and methyl jasmonate-mediated changes in flavonoid concentrations and antioxidant activity in broccoli florets and kale leaf tissues. *Hort. Science.* 48, 996–1002.

Liang, Y. S., Choi, Y. H., Kim, H. K., Linthorst, H. J. M., and Verpoorte, R. (2006). Metabolomic analysis of methyl jasmonate treated *Brassica* rapa leaves by 2-dimensional NMR spectroscopy. *Phytochemistry* 67, 2503–2511. doi: 10.1016/j.phytochem.2006.08.018

Liu, A. G., Juvik, J. A., Jeffery, E. H., Berman-Booty, L. D., Clinton, S. K., and Erdman, J. W. (2014). Enhancement of broccoli indole glucosinolates by methyl jasmonate treatment and effects on prostate carcinogenesis. *J. Med. Food.* 17, 1177–1182. doi: 10.1089/jmf.2013.0145

Luthria, D. L. (2012). Optimization of extraction of phenolic acids from a vegetable waste product using a pressurized liquid extractor. *J. Funct. Foods* 4, 842–850. doi: 10.1016/j.jff.2012.06.001

Miglio, C., Chiavaro, E., Visconti, A., Fogliano, V., and Pellegrini, N. (2008). Effects of different cooking methods on nutritional and physicochemical characteristics of selected vegetables. *J. Agric. Food Chem.* 56, 139–147. doi: 10.1021/jf072304b

Mikkelsen, M. D., Larsen-Petersen, B., Glawischnig, E., Bøgh-Jensen, A., Andreasson, E., and Halkier, B. A. (2003). Modulation of CYP79 genes and glucosinolate profiles in arabidopsis by defense signaling pathways. *Plant Physiol.* 131, 298–308. doi: 10.1104/pp.011015

Natella, F., Nardini, M., Di Felice, M., and Scaccini, C. (1999). Benzoic and cinnamic acid derivatives as antioxidants: structure – activity relation. *J. Agric. Food Chem.* 47, 1453–1459. doi: 10.1021/jf980737w

Pérez-Balibrea, S., Moreno, D. A., and García-Viguera, C. (2011). Improving the phytochemical composition of broccoli sprouts by elicitation. *Food Chem.* 129, 35–44. doi: 10.1016/j.foodchem.2011.03.049

Porras-Baclayon, D., Matsui, T., Haruo, S., and Kosugi, Y. (2007). Changes in activity and gene expression of phenylalanine ammonia-lyase in broccoli branchlets during early postharvest senescence. *J. Biol. Sci.* 7, 614–619. doi: 10.3923/jbs.2007.614.619

Radojčić-Redovniković, I., Glivetić, T., Delonga, K., and Vorkapić-Furač, J. (2008). Glucosinolates and their potential role in plant. *Period. Biol.* 110, 297–309.

Reyes, L. F., and Cisneros-Zevallos, L. (2003). Wounding stress increases the phenolic content and antioxidant capacity of purple-flesh potatoes (*Solanum tuberosum* L.). *J. Agric. Food Chem.* 51, 5296–5300. doi: 10.1021/jf034213u

Rice-Evans, C., Miller, N., and Paganga, G. (1997). Antioxidant properties of phenolic compounds. *Trends Plant Sci.* 2, 152–159. doi: 10.1016/S1360-1385(97)01018-2

Rodríguez-Cantú, L. N., Gutiérrez-Uribe, J. A., Arriola-Vucovich, J., Díaz-De La Garza, R. I., Fahey, J. W., and Serna-Saldivar, S. O. (2011). Broccoli (*Brassica oleracea* var. italica) sprouts and extracts rich in glucosinolates and isothiocyanates affect cholesterol metabolism and genes involved in lipid homeostasis in hamsters. *J. Agric. Food Chem.* 59, 1095–1103. doi: 10.1021/jf103513w

Saha, S., Hollands, W., Teucher, B., Needs, P. W., Narbad, A., Ortori, C. A., et al. (2012). Isothiocyanate concentrations and interconversion of sulforaphane to erucin in human subjects after consumptions of commercial frozen broccoli compared to fresh broccoli. *Mol. Nutr. Food Res.* 56, 1906–1916. doi: 10.1002/mnfr.201200225

Starzyńska, A., Leja, M., and Mareczek, A. (2003). Physiological changes in the antioxidant system of broccoli flower buds senescing during short-term storage, related to temperature and packaging. *Plant Sci.* 165, 1387–1395. doi: 10.1016/j.plantsci.2003.07.004

Surjadinata, B. B., and Cisneros-Zevallos, L. (2012). Biosynthesis of phenolic antioxidants in carrot tissue increases with wounding intensity. *Food Chem.* 134, 615–624. doi: 10.1016/j.foodchem.2012.01.097

Tian, M. S., Downs, C. G., Lill, R. E., and King, G. A. (1994). A role of ethylene in the yellowing of broccoli after harvest. *J. Am. Soc. Hortic. Sci.* 119, 276–281.

Torres-Contreras, A. M., Nair, V., Cisneros-Zevallos, L., and Jacobo-Velázquez, D. A. (2014a). Plants as biofactories: stress-induced production of chlorogenic acid isomers in potato tubers as affected by wounding intensity and storage time. *Ind. Crop Prod.* 62, 61–66. doi: 10.1016/j.indcrop.2014.08.018

Torres-Contreras, A. M., Nair, V., Cisneros-Zevallos, L., and Jacobo-Velázquez, D. A. (2014b). Effect of exogenous amylolytic enzymes on the accumulation of chlorogenic acid isomers in wounded potato tubers. *J. Agric. Food Chem.* 62, 7671–7675. doi: 10.1021/jf5026983

Vallejo, F., Tomás-Barberán, F., and García-Viguera, C. (2003). Health-promoting compounds in broccoli as influenced by refrigerated transport and retail sale period. *J. Agric. Food Chem.* 51, 3029–3034. doi: 10.1021/jf021065j

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153, 895–905. doi: 10.1104/pp.110.155119

Verkerk, R., Dekker, M., and Jongen, W. M. F. (2001). Post-harvest increase of indolyl glucosinolates in response to chopping and storage of *Brassica* vegetables. *J. Sci. Food Agric.* 81, 953–958. doi: 10.1002/jsfa.854

Vinson, J. A., Hao, Y., Su, X., and Zubik, L. (1998). Phenol antioxidant quantity and quality in foods: vegetables. *J. Agric. Food Chem.* 46, 3630–3634. doi: 10.1021/jf980295o

Whetten, R., and Sederoff, R. (1995). Lignin biosynthesis. *Plant Cell* 7, 1001–1013. doi: 10.1105/tpc.7.7.1001

Yuan, G., Sun, B., Yuan, J., and Wang, Q. (2010). Effect of 1-methylcyclopropene on shelf life, visual quality, antioxidant enzimes and health-promoting compounds in broccoli florets. *Food Chem.* 118, 774–781. doi: 10.1016/j.foodchem.2009.05.062

Zhang, L., Wang, G., Chang, J., Liu, J., Cai, J., Rao, X., et al. (2010). Effects of 1-MCP and ethylene on expression of three CAD genes and lignification in stems of harvested Tsai Tai (*Brassica* chinensis). *Food Chem.* 123, 32–40. doi: 10.1016/j.foodchem.2010.03.122

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Villarreal-García, Nair, Cisneros-Zevallos and Jacobo-Velázquez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 February 2016
doi: 10.3389/fpls.2016.00021

Characterizing Variation of Branch Angle and Genome-Wide Association Mapping in Rapeseed (*Brassica napus* L.)

*Jia Liu[1], Wenxiang Wang[1], Desheng Mei[1], Hui Wang[1], Li Fu[1], Daoming Liu[2], Yunchang Li[1] and Qiong Hu[1]**

[1] Key Laboratory of Biology and Genetic Improvement of Oil Crops, Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy of Agricultural Sciences, Wuhan, China, [2] Agricultural Sciences Institute of Lu'an Municipal, Lu'an, China

Changes in the rapeseed branch angle alter plant architecture, allowing more efficient light capture as planting density increases. In this study, a natural population of rapeseed was grown in three environments and evaluated for branch angle trait to characterize their phenotypic patterns and genotype with a 60K *Brassica* Infinium SNP array. Significant phenotypic variation was observed from 20 to 70°. As a result, 25 significant quantitative trait loci (QTL) associated with branch angle were identified on chromosomes A2, A3, A7, C3, C5, and C7 by the MLM model in TASSEL 4.0. Orthologs of the functional candidate genes involved in branch angle were identified. Among the key QTL, the peak SNPs were close to the key orthologous genes *BnaA.Lazy1* and *BnaC.Lazy1* on A3 and C3 homologous genome blocks. With the exception of Lazy (*LA*) orthologous genes, *SQUMOSA PROMOTER BINDING PROTEIN LIKE 14* (*SPL14*) and an auxin-responsive *GRETCHEN HAGEN 3* (*GH3*) genes from *Arabidopsis thaliana* were identified close to two clusters of SNPs on the A7 and C7 chromosomes. These findings on multiple novel loci and candidate genes of branch angle will be useful for further understanding and genetic improvement of plant architecture in rapeseed.

Keywords: *Brassica napus* L., branch angle, genetic variation, association mapping, multiple environments, candidate genes

OPEN ACCESS

Edited by:
Donal Martin O'Sullivan, University of Reading, UK

Reviewed by:
Maoteng Li, Huazhong University of Science and Technology, China
Daniela Marone, Centre of Cereal Research, Italy

***Correspondence:**
Qiong Hu
huqiong01@caas.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *01 October 2015*
Accepted: *08 January 2016*
Published: *04 February 2016*

Citation:
Liu J, Wang W, Mei D, Wang H, Fu L, Liu D, Li Y and Hu Q (2016) Characterizing Variation of Branch Angle and Genome-Wide Association Mapping in Rapeseed (Brassica napus L.). Front. Plant Sci. 7:21. doi: 10.3389/fpls.2016.00021

INTRODUCTION

Rapeseed (*Brassica napus* L. 2n = 4 × = 38, AACC genomes) is a widely cultivated oil crop throughout the world. Yield improvement and mechanized harvesting are extremely urgent recently for the demands of rapeseed producers in addition to the edible oil and biofuel industries (Diepenbrock, 2000). The ideotype of a plant is defined as the spatial distribution of various architectures which is an important agronomic character that affects photosynthesis and seed yields (Donald, 1968; Mansfield and Mumm, 2014). The ideotype can influence photosynthesis, plant growth, and seed yield due to the least competition among the individuals in a population (Wang and Li, 2005). The ideotype is determined by a combination of architecture factors including branch angle (BA), plant height (PH), first branch height (FBH), inflorescence length (IL), and branch number (BN; Mei et al., 2009; Shi et al., 2009; Xu et al., 2014). In particular, branch angle, or the angle between branch and erect primary stem, has long attracted the attention of breeders because of the significant contribution of this trait to plant architecture.

High yields can be achieved through high plant density with a small branch angle, which determines the plant's ability to grow and capture light efficiently. Wang and Li (2008) reported in rice more upright and dense leaves not only improve light capture but also improve the accumulation of leaf nitrogen for grain filling. The branch (or leave) angle has also been studied in maize, cotton, and other important crops to achieve an ideal plant architecture to improve yields (Li et al., 2007; Jin et al., 2008; Song and Zhang, 2009; Ku et al., 2011; Tian et al., 2011; Bai et al., 2012). Several quantitative trait loci (QTL) related to leaf angle have been identified genetically in maize.

In recent years, genome-wide association (GWA) mapping has become a powerful tool for identifying important genes associated with complex traits, which has been used with success in model and non-model plants (Breseghello and Sorrells, 2006; Atwell et al., 2010; Huang et al., 2010; Zhao et al., 2011). Additionally, with the dramatically decreasing cost of genome sequencing and rapid developments in genome analysis, *Brassica* A genome sequence from *Brassica rapa* and *Brassica* C genome sequence from *Brassica oleracea* have been published (Wang et al., 2011; Liu et al., 2014). It is important to complete the rapeseed (*B. napus*, AC genome) genome sequencing (Chalhoub et al., 2014). Single nucleotide polymorphisms (SNPs) are abundant and evenly distributed throughout the genomes, which is more conducive to GWA. In 2012, a *B. napus* 60K SNP Infinium genotyping array was produced and applied by the international *Brassica* SNP consortium in cooperation with Illumina Inc. San Diego, CA, USA (Snowdon and Iniguez Luy, 2012; Edwards et al., 2013). Studies about GWAS in rapeseed have gained attention in recent years and various traits including flowering time, seed weight and seed quality have been dissected (Harper et al., 2012; Cai et al., 2014; Li et al., 2014; Lu et al., 2014; Raman et al., 2014; Wang et al., 2014). However, no report was found on association mapping for rapeseed branch angle to our knowledge.

Numerous genes are known to influence the branch (tiller) angle in model and other crops. For instance, *OsTAC1* was reported to play a critical role in controlling rice architecture (Yu et al., 2007). Li et al. (2007) reported that *LAZY1* regulates shoot gravitropism through which the rice tiller angle is controlled. Enlarged branch angles with agravitropic shoots were similarly also found in *Atlazy1* mutant of *Arabidopsis* (Yoshihara et al., 2013). Dong et al. (2013) identified maize *ZmLA1* gene as a functional ortholog of *LAZY1* in rice and *Arabidopsis*. The regulation of branch angle is a combination of environmental factors and hormone homeostasis (Lomax, 1997). Auxin may be the primary hormone involved in shoot gravitropism (Robert and Friml, 2009), which is a key process in determining branch (leave) angle. The member of *GH3* family plays crucial roles in auxin homeostasis in relation to leaf inclination control (Zhao et al., 2012). Recent evidence showed that new strigolactone plant hormones regulated rice tiller angle by attenuating shoot gravitropism through the inhibition of auxin biosynthesis (Sang et al., 2014). Although interaction of auxin and strigolactone plays an important role in shoot gravitropism, the key genes by which the two hormones regulate shoot gravitropism is not yet identified. However, knowledge of the genes that control genetic variation for branch angle of rapeseed is limited.

In this study, a panel of 143 elite rapeseed accessions was analyzed by the 60 K *Brassica* Infinium SNP array. The branch angle was measured in 2013–2014 at three environments. The SNPs in the array were *in silico* mapped to A and C genomes of Darmor-*bzh B. napus* genome "pseudomolecules" to obtain their hypothetical position. The aims were (1) to gain the population structure and genetic diversity in elite germplasms; (2) to detect QTL controlling branch angle and mine for elite alleles; and (3) to predict the candidate genes.

METHODS

Plant Material and Field Experiments

A total of 143 rapeseed accessions were used for an association analysis in this study. According to the information from field observations, the accessions were classified to three different germplasm types, i.e., spring oilseed rape (OSR) (13), semi-winter OSR (124), and winter OSR (6). Based on their origins, 112 accessions originated from China, 24 from Oceania, 5 from Europe, 1 from North America, and 1 from India (Supplementary Table 1). The seeds from all the accessions were collected, stored and supplied by Oil Crops Research Institute of Chinese Academy of Agricultural Sciences (OCRI-CAAS). In recent decades, these accessions have been widely used as parents in breeding programs.

The experiments were conducted at Yangluo Agronomic Experimental Station of OCRI-CAAS (28°42′N, 112°33′E), Wuhan, China, during the 2013 and 2014 winter growing season and at Lu'an Experimental Station (31°73′N, 116°52′E) in Anhui, China, during the 2013 winter growing season. In each environment, the experiment was conducted in randomized complete blocks with two replicates. Each plot contained three rows with 54 individuals, setting as 33 cm between rows and 11 cm between plants within each row, with a planting density of 270,000 plants/ha. All experiments were performed under local field management and cultivation conditions.

Trait Measurements and Statistical Analysis

In each plot, five typical plants were harvested for branch angle measurement at the mature stage. The branch angle (BA) was measured as the angle between the main stem and the branch. BA was scanned with a digital camera (SONY DSLR-A350; SONY, Japan). The angle value from the images was obtained using AutoCAD software (Autodesk Inc., San Rafael, CA). The average BA value of five individual plants for each plot was calculated as the final phenotypic value. In addition, the best linear unbiased estimators (BLUPs) across all three environments were predicted by assuming fixed genotypic effects to minimize the effects of environmental variation. Finally, each environment and BLUP value was used as a phenotype for the association analysis.

Statistical analysis of the data was performed by using PROC MEANS in SAS software, Version 9.3 (2000, SAS Institute Inc., Cary, NC, US). Analysis of variance (ANOVA) was conducted

by using PROC GLM to determine the effects of block, environment, genotype and genotype-environment interactions. Correlation coefficients were obtained by using PROC CORR. The Broad-sense heritability (H^2) was calculated as $H^2 = \sigma g^2/(\sigma g^2 + \sigma gl^2/n + \sigma e^2/nr)$, in which σg^2, σgl^2, σe^2, r, and n represent the estimated variances for the genetic effects, genotype-environment interactions, random errors, number of replications and number of environments, respectively. The estimated variances for σg^2, σgl^2, and σe^2 were obtained by ANOVA.

SNP Genotyping, Quality Control and *In silico* Mapping of SNPs

The genotype of 143 accessions was detected with a *Brassica* 60K Illumina Infinium SNP array according to the work flow by Emei Tongde Co. All of the SNP data were clustered and called up automatically by using Illumina BeadStudio genotyping software. The SNP quality was checked and comparable with previous studies (Li et al., 2014; Wang et al., 2014). The low quality SNP loci (call rate < 80% and/or minor allele frequency < 0.05) in all accessions were deleted from the results. Out of 52,157 SNPs in the array, 2836 that had a zero call frequency of AA or BB were excluded according to the quality control. Using a cut-off for missing data >0.2 and MAF < 0.05, 1860 and 1909 SNPs were filtered, respectively, reducing the number of SNPs to 38,063.

SNP mapping was performed as previously reported (Altschul et al., 1990). In brief, BLAST search against the "pseudomolecules" representative of the *B. napus* genome (version 4). Only the top BLAST hits with an *E*-value cut-off of 1*E*-15 against the pseudomolecules were retained, while BLAST matches to multiple loci with the same *E*-value were deleted. A final set of 34,469 high-performing SNPs was used for all the analyses.

Population Genetic Analysis

Nei's genetic distance matrix for all SNPs is the distance and it was calculated to build unrooted neighbor-joining trees by using PowerMarker (Liu and Muse, 2005). The result was visualized by using FigTree software based on 34,469 SNPs (http://tree.bio.ed.ac.uk/software/figtree/). Kinship (K) matrix used to compare all the pairs of the 143 accessions was calculated with 2434 informative SNPs with a MAF > 0.2 by using the SPAGeDi software package (Hardy and Vekemans, 2002). All negative kinship values that were found between two individuals, which indicates that there was less of a relation than expected between two random individuals, were transformed to zero (Yu et al., 2005).

A total of 2434 SNPs [minor allele frequency (MAF) ≥ 0.2] that were evenly distributed across the whole genome were selected to create the population structure inferred by using the STRUCTURE v2.3.4 software package (Pritchard et al., 2000). Iterations were performed 100,000 times by using a burn-in length of 100,000 MCMC (Markov chain Monte Carlo) with the admixture and related frequency model. Five independent runs were performed with *K*-values (the putative number of populations) ranging from 1 to 10. The optimal *K*-value was determined by taking the log probability of the data [LnP(D)] and an *ad hoc* statistical Δk based on the rate of change for LnP(D) between successive *k*-values as described by Evanno et al. (2005). The cluster membership coefficient matrices of replicate runs from STRUCTURE were integrated to obtain a Q matrix by using CLUMPP software (Jakobsson and Rosenberg, 2007) and graphically displayed by using the DISTRUCT software package (Rosenberg, 2004). Accessions with a probability of membership >0.7 were assigned to corresponding clusters, and those <0.7 were assigned to a mixed group. The population structure matrix (Q) was generated for further analyses. Linkage disequilibrium (LD) parameter (r^2) for estimating the degree of LD between pair-wise SNPs (MAF ≥ 0.05) was calculated using the software TASSEL 4.0 (Bradbury et al., 2007).

Association Analysis

Two different models were used to test associations. The first model was a simple and general linear model (GLM) without controlling for Q and K, containing only the SNP that was tested as a fixed effect. The second model was a mixed linear model (MLM) where, in addition to testing the SNP, the population structure (Q) and relative kinship matrix were included as fixed and random effects, respectively. Analyses were performed by using TASSEL 4.0 software, for which the optimum compression and population parameters previously determined (P3D) variance component estimation were implemented to decrease the computing time for the large data set (Zhang et al., 2010).

The significance of associations between SNPs and trait was based on a threshold of $p < 2.90 \times 10^{-5}$ [i.e., $-\log_{10}(p) = 4.5$]. The threshold is 2.90×10^{-5} at a significant level of 1% after Bonferroni multiple test correction (1/34,496). Furthermore, we applied the false discovery rate (FDR) technique. We calculated an FDR *q*-value for each association test by using the software QVALUE (Dabney and Storey, 2004). The FDR *q*-value of the significant SNP with the lowest test statistic ($P < 0.05$) provided an estimate of the proportion of false positives among the significant associations. The significant value and the marker effect for each SNP were exported, and a Manhattan plot was generated in the R package qqman (Turner, 2014, http://cran.r-project.org/web/packages/qqman/).

Stepwise regression was performed to detect the effect of multiple alleles with different functional polymorphisms on branch angle and to estimate the total variance explained (R^2) by using the *lm* function in R (Ihaka and Gentleman, 1996).

RESULTS

Phenotypic Variation of Branch Angle

Significant variation was observed among the 143 rapeseed accessions in the three environments studied, with branch angles ranging from 20 to 70° (**Figure 1**, **Table 1**). In three environments Yangluo-2013, Yangluo-2014, and Lu'an-2013, the natural population exhibited average branch angle values (±SD) of 39.93 ± 5.98, 37.52 ± 7.26, and 42.30 ± 6.65, respectively. The frequency distributions of branch angle in the natural population are summarized in **Figure 2**. Branch angle ranged from 22.32 to

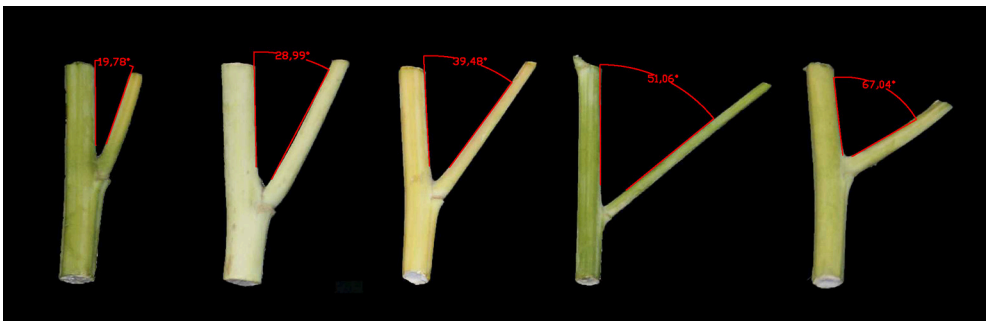


FIGURE 1 | Branch angle was measured using AutoCAD software.

TABLE 1 | Phenotypic characteristics for branch angle in 143 rapeseed accessions.

Environment	Min	Max	Mean ± SD	CV%	Kurtosis	Skewness	Correlation coefficient		
							13YL	14YL	13LA
YL-2013	22.32	59.57	39.93 ± 5.99	14.99	0.91	0.08		0.65**	0.68**
YL-2014	20.04	69.09	37.52 ± 7.26	19.34	1.92	0.31			0.56**
LA-2013	22.98	70.77	42.30 ± 6.65	15.71	2.22	0.45			

SD, standard deviation; CV, coefficient of variation; Kurtosis is the distribution of observed data around the mean; Skewness is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean.

***Significant at $P < 0.01$.*

YL, Yangluo; LA, Lu'an.

59.57, 20.04 to 69.09, and 22.98 to 70.77 in three environments, respectively.

The two-way ANOVA showed that differences among the lines in the branch angle were highly significant. This finding confirmed that a large number of genetic variation existed in the population. The effects from years and location were significant (**Table 2**). The broad-sense heritability (H^2) for branch angle was calculated as 93.90%, demonstrating that branch angle in these rapeseed lines was conditioned primarily by genetic factors. The correlation coefficients for branch angle between environments were all relatively high ($r \geq 0.557$, $P < 0.01$).This observation indicated that branch angle in the 143 lines was relatively consistent across environments.

SNP Genetic Diversity and Linkage Disequilibrium

Supplementary Table 2 gives the information on all of the SNPs. Estimates of an average nucleotide diversity (polymorphism information content or PIC) of 0.366 showed that the overall genetic variation in the accessions studied here represents ∼62.9% of the rapeseed diversity. There had the most markers (2745) with a marker density of one per 26.4 kb in C3 linkage group and the fewest markers (932) with a marker density of one per 47.6 kb in C5.

To evaluate the extent of LD, r^2 was used to calculate LD. The genome-wide LD decay of A and C genome for rapeseed germplasms are shown in **Figure 3**. Taken together, the LD of A genome decayed significantly faster than that of the C genome. We estimated the LD decay, r^2 decayed to 0.2 when the average distance for A genome was 206 kb and r^2 decayed to 0.2 when the average distance for C genome was 949 kb.

Population Structure and Relative Kinship

Population structure of the association panel was calculated using 2434 SNPs, and a clustering inference performed with possible clusters (k) from 1 to 10 showed that the most significant change in likelihood occurred when K increased from 2 to 3, and the highest Δk-value was observed at $k = 2$ (**Figure 4**). Based on the Δk method (**Figure 4B**), the 143 accessions could be divided into two sub-populations (**Figure 4C**). By using a probability-of-membership threshold of 70%, 99 and 17 lines were assigned to the two groups, respectively. The remaining 27 lines were classified into a mixed group (Supplementary Table 1). In comparison with a previous study, this population structure classification yields the same results (Harper et al., 2012; Lu et al., 2014; Wang et al., 2014). In addition, the NJ phylogenetic tree based on Nei's genetic distances displayed two clear clades (Supplementary Figure 1), corresponding to the two groups estimated by STRUCTURE. The lines belong to mixed groups were distributed across the whole tree. Tree-based analyses yielded results very similar to those of the STRUCTURE analysis.

The 2434 informative SNPs with a MAF > 0.2 and little or no missing data were used to estimate the relative kinship in the set of 143 lines. As shown in **Figure 5**, the average relative kinship between any two lines was 0.0332, or ∼57% of the pairwise kinship estimates were close to 0, and 21% of the kinship estimates ranged from 0 to 0.05. The remaining estimates ranged

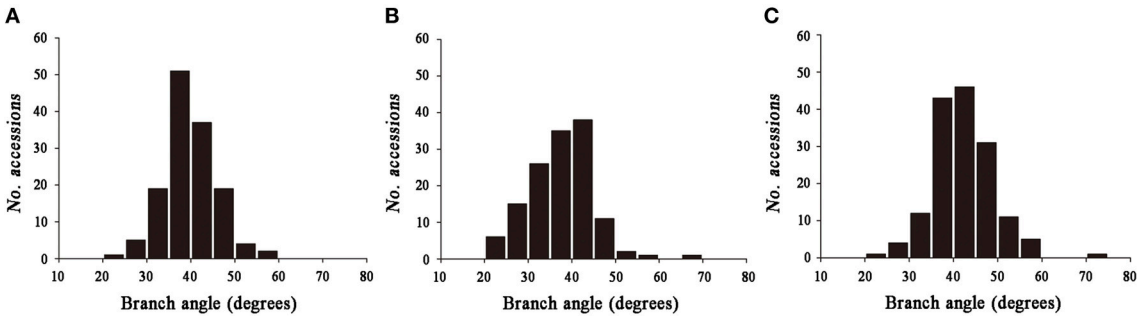


FIGURE 2 | The frequency distribution for branch angle in three environments. (A) 2013YL; (B) 2014YL; (C) 2013LA.

TABLE 2 | ANOVA and broad-sense heritability for branch angle.

Variance resource	DF	SS	MS	*F*-value	H^2(%)
Environments	2	1210.86	605.43	139.56**	93.90
Gene	142	18487.16	133.96	30.88**	
Rep	1	20.01	20.01	4.61*	
G×E	305	6513.11	23.95	5.52**	

*DF, degree freedom; SS; stdev square; MS, mean square; F-value, F-test value; H^2, heritability; G×E, genotype by environment interactions; *, ** Significant at $P < 0.05$ and 0.01 probability levels, respectively.*

from 0.05 to 1, with a continuously decreasing number of pairs falling in higher estimated categories. These results indicated that most lines in the panel have no or very weak kinship, which might be attributed to the broad ranging collection of genotypes and the exclusion of similar genotypes before analysis.

Association Mapping and Candidate Gene Prediction

A total of 25 and 60 associations ($P < 2.90E^{-5}$) were detected for branch angle by using BLUP across three environments and an individual environment (**Figure 6**, **Table 3**; Supplementary Figure 2). To select the major QTL among all the significant SNPs, these SNPs were clumped by using an LD block as a criterion (Gabriel et al., 2002), and the peak SNP within each LD block was retained. After the clumping of SNPs, six QTL for branch angle were distinguished with the BLUP values of a branch angle across three environments and the peak SNPs are listed in **Table 3**. The six peak regions were 6.1, 3.5, 23.2, 3.0, 39.5, and 48.7 Mb of A2, A3, A7, C3, C5, and C7, respectively, in the "pseudomolecules" of *B. napus*, and the cumulative phenotypic variance explained by all significant SNPs was 92.42%, which contributed to 16.60–18.95% of the phenotypic variance based on the R^2-values. In Yangluo-2013, nine peak SNPs were detected in Q+K models with an FDR of 0.021, and the eight peak regions were 10.1, 22.2, 18.0, 0.1, 42.8, 48.7, and 2.9 Mb for A2, A5, A8, A10, C2, C6, C7, and C8, respectively. In Yangluo-2014, two peak SNPs were detected with FDRs of 0.013 and 0.027, and the two peak regions were 48.7 and 23.2 Mb of A7 and C7, respectively. In Lu'an-2013, only one significantly peak SNP in A3 was detected in 31.3 Mb region with an FDR of 0.263.

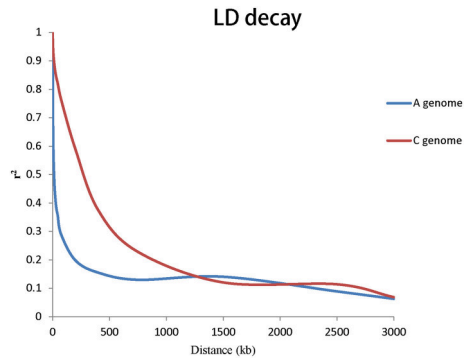


FIGURE 3 | Genome-wide LD decay of A and C genomes for all the 143 accessions.

Notably, the *Lazy* orthologs were searched in the "pseudomolecules" of *B. napus* and two orthologs were found at 2.6 Mb in A3 and 3.2 Mb in C3, which were 488 kb away from the peak SNP Bn-A03-p3571859 and 243 kb away from the peak SNP Bn-A03-p3571859. In addition, the *Squamosa Promoter Binding Protein-like* 14 (*SPL14*) orthologs were searched in the "pseudomolecules" of *B. napus* and one ortholog was found at 24.3 Mb of A7, which was 521 kb away from the peak SNP Bn-A07-p19977445. Aside from the *Lazy* and *SPL14* genes, an auxin-responsive *GRETCHEN HAGEN 3* (*GH3*) ortholog "pseudomolecules" of *B. napus* in 48.7 Mb of C7, which was very close (only 4 kb) to the peak SNP Bn-scaff-16110-1-p1940585.

DISCUSSION

In this study, a panel comprising 143 *B. napus* germplasm lines was adopted for association mapping study. Significant natural phenotypic variation in branch angle was observed in rapeseed germplasms. In general, to unravel the genetic basis of trait variation, a number of traditional genome-wide phenotype-to-genotype approaches have been employed by LD association mapping. An association mapping study employs the large number of recombination events that occur throughout the entire breeding selection history of the mapping population, thereby allowing fine-scale QTL mapping (Nordborg and Tavaré,

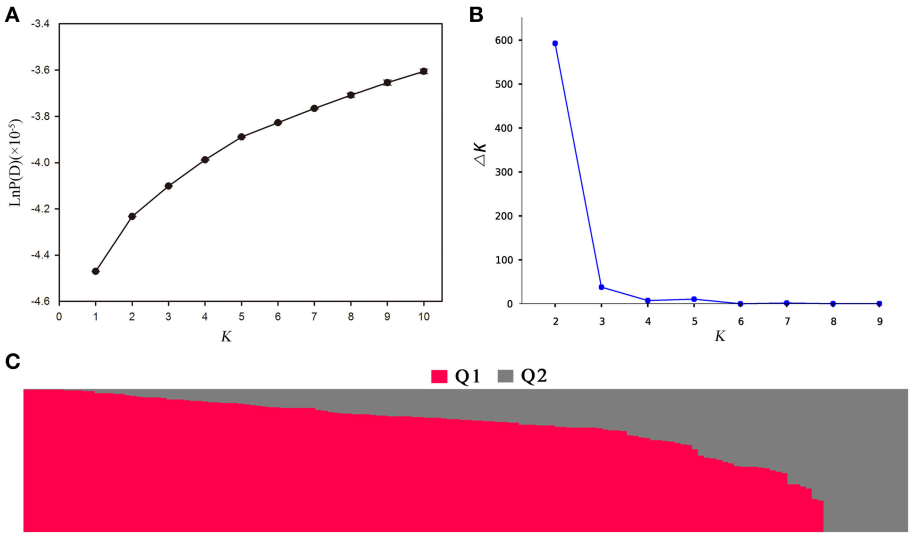


FIGURE 4 | Population structure analysis of 143 rapeseed accessions by STRUCTURE software. (A) The estimated LnP(D) of possible clusters (k) from 1 to 10; **(B)** Δk based on the change of LnP(D) between consecutive k; and **(C)** Q1 and Q2 are the composition values belonging to the two sub-populations ($K = 2$) for a given accession which is represented by a vertical bar.

2002). Although the sample size is not sufficiently large in our association panel, the phenotypic variation in branch angle is very large. The heritability of this trait is relatively high, and it is related to significant genome loci with great effects. However, there must be existence of some unknown genes regulating branch angle in rapeseed comparing to model plant *Arabidopsis* and rice (Teichmann and Muhr, 2015), which still could not make of it. The novelty SNP clusters of QTL from our preliminary exploration are starting to decompose this aspect. Based on the MLM model, a total of 60 SNP associations ($P < 2.90E^{-5}$) were detected for branch angle in three environments, and 25 significant SNP loci were further verified using BLUP model. Through the analysis of GWA, the markers detected in the environmental BLUPs value correlated significantly ($P < 0.001$) with branch angle, with phenotypic value effects between 16.60 and 18.93%. Exploring these associated markers provides a genetic basis to analyze branch angle variation in rapeseed.

The discovery of many false-positive QTL is due to the population structure (Zhao et al., 2007). To resolve this problem, several models have been developed including the Q+K model and the PCA model. Although the Q+K model have been demonstrated the most powerful method for identifying associations by many studies (Yu et al., 2005; Stich and Melchinger, 2009), we also compared different models and obtained the same conclusion that the MLM model (Q+K) was most suitable for our population. In our study, all *B. napus* accession lines can be largely divided into two sub-populations in STRUCTURE and compared with an association analysis from a previous study (Harper et al., 2012), and the population structure classification leads to the same results. Hence, the reasonable results for a population structure-associated analysis provide a foundation and guarantee. For the association analysis, we used a mixed model approach that avoided a confounding effect in the population structure and population relations. We acquired the branch angle BLUP values in three independent environments to eliminate the influence of environment.

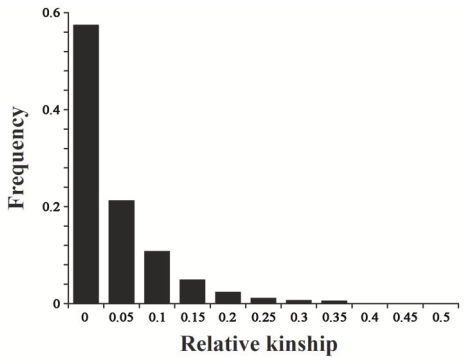


FIGURE 5 | The distribution for pairwise relative kinship. Only the kinship values ranging from 0 to 0.5 are shown.

To account for the potential number of false positives, we implemented stringent quality control on the included polymorphisms, with conservative non-parametric testing and an adjusted statistical significance threshold. Our approach relied on natural variations in rapeseed, leading to a set of strong candidate genes by comparative genome method. A high degree of co-linearity and congruence in the *A. thaliana* and *Brassica* genomes has been recognized (Parkin et al., 2005; Ziolkowski et al., 2006) and a single locus in the *Arabidopsis* genome is generally represented by three distinct loci in diploid *Brassica* species. Several genes that play key roles in regulating branch development in *Arabidopsis* have been identified. By

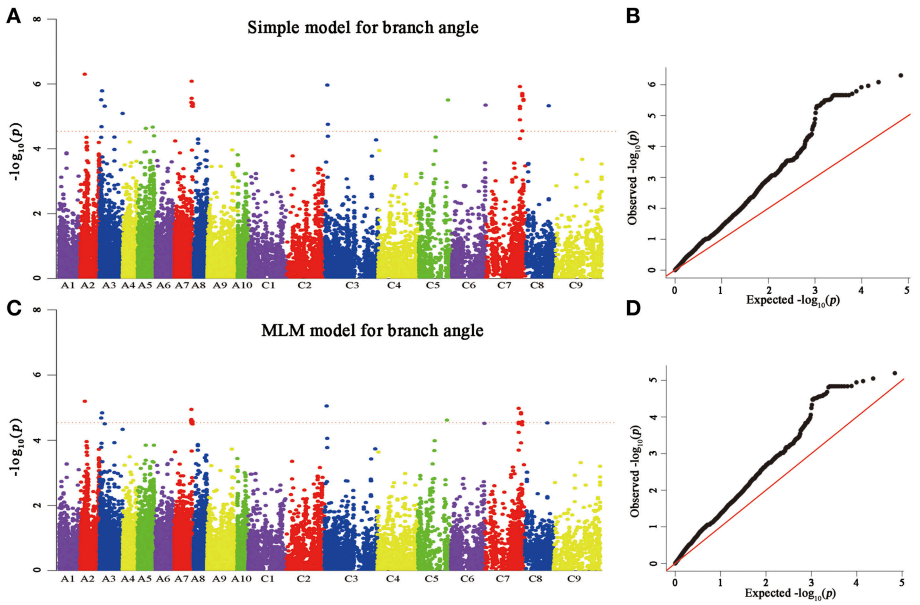


FIGURE 6 | Genome-wide association analysis of branch angle. (A) Manhattan plots of the simple model for BA. **(B)** A quantile-quantile (QQ) plot for a simple branch model **(C)** Manhattan plots of MLM for BA, as in **(A)**. **(D)** A quantile-quantile (QQ) plot of MLM for BA. Each dot represents a SNP. The horizontal dashed red line indicates the Bonferroni-corrected significance threshold at $-\log_{10}(P) = 4.5$.

applying this method to the recently grown tetraploid crop *B. napus*, we identified genomic regions that underlie four QTL for branch angle. The detected regions contained two *Lazy* (AT5G14090) orthologous, *SPL14* (AT1G20980) and Auxin-responsive *GH3* family protein (AT5G51470) genes, which were all relevant to angle formation in *Arabidopsis thaliana*.

Rapeseed is an excellent model for association analyses because of the extensive architectural variation across its native range and its diverse germplasm collections through artificial selection. During rapeseed domestication and breeding process, breeders often focus on important characteristics such as oil content, biotic stress resistance, and quality for directional selection for greater oil yield and better quality, which were directly caused by an artificial selection gene locus mutation at a lower frequency. For example, in terms of oil and glucosinolate contents, and because of the artificial selection on *Fatty Acid Elongation 1* (*FAE1*) and *High Aliphatic Glucosinolate 1* (*HAG1*) gene locus, gene locus mutation occurs at a lower frequency in double-low rapeseed populations (Li et al., 2014). However, the optimation of rapeseed architecture (branch angle and plant height) to improve yield is more urgent nowadays. Association-mapping approach has advantage for distinguishing the most favorable alleles within a diverse genetic background, which provides the necessary genotypic information to facilitate the design of efficient rapeseed introgression and selection schemes throughout the world. This study indicated that the SNP variation frequency control of rapeseed branch angle habits is slightly higher than the results. This finding may be related to the domestication and genetic improvement of characteristics that are not particularly selected by the breeding program.

The association of branch angle with increased plant productivity in rapeseed may be beneficial to lodging resistance and high-density field cultivation. This genetic analysis of branch angle and plant morphological characteristics was only based on one natural population. Therefore, the observed inter-trait and trait-marker associations must be confirmed in other rapeseed populations. Further evaluations of branch angle and plant morphology should be conducted under contrasting cropping density, and total seed yield measurements should be taken because the seed harvest index in addition to the plant growth and productivity are strongly affected by the cropping density.

CONCLUSION

In this study, GWA mapping with corrections for the population structure were used to identify a number of novel loci and refine the map locations of known loci related to branch angle in rapeseed. This information not only demonstrates that GWAS mapping can be used as a powerful tool for dissecting plant architecture mechanisms in rapeseed but also provides valuable markers for breeding rapeseed cultivars with the ideotype. In addition, the candidate genes nearby these SNP loci represent promising targets for efforts to further identify causal variants and to clarify how the implicated genes affect branch angle in rapeseed.

AUTHOR CONTRIBUTIONS

JL and QH conceived and designed the study. JL, WW, and DL conducted the experiments; DM and HW coordinated

TABLE 3 | A summary of significant ($P < 2.90E^{-5}$) SNP-trait associations associated with branch angle.

Environment	Peak SNP name	Chr[a]	Position[b] (kb)	*P*	MAF[C]	R^2 (%)	Candidate gene	Marker distance from candidate gene (kb)[d]
BA-BLUPS	Bn-A02-p7183200	A2	6140	6.38E-06	(A)0.084	18.93		
	Bn-A03-p3571859	A3	3571	1.46E-05	(G)0.310	17.53	*BnaA.Lazy.A3*	488 (Upstream)
	Bn-A07-p19977445	A7	23,182	1.14E-05	(G)0.196	18.58	*BnaA.SPL14.A7*	521 (Downstream)
	Bn-scaff_18936_1-p210783	C3	2962	8.92E-06	(A)0.286	18.41	*BnaC.Lazy.C3*	243 (Downstream)
	Bn-scaff_21369_1-p378189	C5	39,469	2.43E-05	(G)0.304	16.60		
	Bn-scaff_16110_1-p1940585	C7	48,687	1.43E-05	(C)0.141	17.56	Auxin-responsive GH3 family protein	4 (Upstream)
BA-13YL	Bn-A02-p11153317	A2	10,098	1.43E-05	(A)0.268	17.57	*BnaA.SPL14.A2*	898 (Downstream)
	Bn-A05-p22938228	A5	22,197	2.61E-05	(A)0.185	16.54		
	Bn-A08-p20838949	A8	18,001	2.75E-06	(C)0.077	20.51		
	Bn-A10-p4727374	A10	137	1.03E-05	(G)0.152	18.14		
	Bn-scaff_15712_2-p770819	C2	42,843	2.10E-05	(A)0.325	17.00		
	Bn-scaff_17799_1-p1178394	C6	45,148	2.61E-05	(G)0.120	18.71	*BnaC.SPL14.C6*	11 (Downstream)
	Bn-scaff_16110_1-p1939090	C7	48,689	2.72E-05	(G)0.129	16.47	Auxin-responsive GH3 family protein	6 (Upstream)
	Bn-scaff_21003_1-p435731	C8	2901	1.90E-05	(A)0.259	17.08		
BA-14YL	Bn-scaff_16110_1-p1940470	C7	48,687	5.43E-06	(A)0.140	19.84	Auxin-responsive GH3 family protein	4 (Upstream)
	Bn-A07-p19977445	A7	23,182	2.15E-05	(G)0.196	18.31		
BA-13LA	Bn-A03-p31327366	A3	31,323	2.90E-05	(A)0.129	16.64		

[a] *Chromosome;*
[b] *The physical position of SNP is inferred from BLAST hits of the chromosome pseudomolecules in B. napus;*
[c] *Minor allele frequency;*
[d] *The marker distance and its upstream or downstream from candidate gene.*

genotyping with SNP markers; LF and YL provided rapeseed lines; JL and WW analyzed and interpreted data, and prepared the manuscript; QH supervised the whole study; all authors reviewed and edited the manuscript.

ACKNOWLEDGMENTS

The Science and Technology Innovation Project of Chinese Academy of Agricultural Sciences (Group No. 118), the Earmarked Fund for China Agriculture Research System (CARS-13), the Hubei Agricultural Science and Technology Innovation Center, the Natural Science Foundation of China (31471535), and Natural Science Foundation of Hubei Province (2014CFB156) supported this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00021

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/s0022-2836(05)80360-2

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., et al. (2010). Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* 465, 627–631. doi: 10.1038/nature08800

Bai, F., Reinheimer, R., Durantini, D., Kellogg, E. A., and Schmidt, R. J. (2012). TCP transcription factor, BRANCH ANGLE DEFECTIVE 1 (BAD1), is required for normal tassel branch angle formation in maize. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12225–12230. doi: 10.1073/pnas.1202439109

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Breseghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172, 1165–1177. doi: 10.1534/genetics.105.044586

Cai, D., Xiao, Y., Yang, W., Ye, W., Wang, B., Younas, M., et al. (2014). Association mapping of six yield-related traits in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 127, 85–96. doi: 10.1007/s00122-013-2203-9

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Dabney, A., and Storey, J. D. (2004). Q-value estimation for false discovery rate control. *Medicine* 344, 539–548.

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): a review. *Field Crop. Res.* 67, 35–49. doi: 10.1016/s0378-4290(00)00082-4

Donald, C. M. (1968). The breeding of crop ideotypes. *Euphytica* 17, 385–403. doi: 10.1007/BF00056241

Dong, Z., Jiang, C., Chen, X., Zhang, T., Ding, L., Song, W., et al. (2013). Maize LAZY1 mediates shoot gravitropism and inflorescence development through regulating auxin transport, auxin signaling, and light response. *Plant Physiol.* 163, 1306–1322. doi: 10.1104/pp.113.227314

Edwards, D., Batley, J., and Snowdon, R. J. (2013). Accessing complex crop genomes with next-generation sequencing. *Theor. Appl. Genet.* 126, 1–11. doi: 10.1007/s00122-012-1964-x

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., et al. (2002). The structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424

Hardy, O. J., and Vekemans, X. (2002). SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes* 2, 618–620. doi: 10.1046/j.1471-8286.2002.00305.x

Harper, A. L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., et al. (2012). Associative transcriptomics of traits in the polyploid crop species *Brassica napus*. *Nat. Biotechnol.* 30, 798–802. doi: 10.1038/nbt.2302

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695

Ihaka, R., and Gentleman, R. R. (1996). A language for data analysis and graphics. *J. Comput. Graph. Statist.* 5, 299–314.

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233

Jin, J., Huang, W., Gao, J. P., Yang, J., Shi, M., Zhu, M. Z., et al. (2008). Genetic control of rice plant architecture under domestication. *Nat. Genet.* 40, 1365–1369. doi: 10.1038/ng.247

Ku, L., Wei, X., Zhang, S., Zhang, J., Guo, S., and Chen, Y. (2011). Cloning and characterization of a putative *TAC1* ortholog associated with leaf angle in maize (*Zea mays* L.). *PLoS ONE* 6:e20621. doi: 10.1371/journal.pone.0020621

Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., et al. (2014). Genome-wide association study dissects the genetic architecture of seed weight and seed quality in rapeseed (*Brassica napus* L.). *DNA Res.* 21, 355–367. doi: 10.1093/dnares/dsu002

Li, P., Wang, Y., Qian, Q., Fu, Z., Wang, M., Zeng, D., et al. (2007). LAZY1 controls rice shoot gravitropism through regulating polar auxin transport. *Cell Res.* 17, 402–410. doi: 10.1038/cr.2007.38

Liu, K., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21, 2128–2129. doi: 10.1093/bioinformatics/bti282

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5:3930. doi: 10.1038/ncomms4930

Lomax, T. L. (1997). Molecular genetic analysis of plant gravitropism. *Gravit. Space Biol. Bull.* 10, 75–82.

Lu, G., Harper, A. L., Trick, M., Morgan, C., Fraser, F., O'Neill, C., et al. (2014). Associative transcriptomics study dissects the genetic architecture of seed glucosinolate content in *Brassica napus*. *DNA Res.* 21, 613–625. doi: 10.1093/dnares/dsu024

Mansfield, B. D., and Mumm, R. H. (2014). Survey of plant density tolerance in US maize germplasm. *Crop Sci.* 54, 157–173. doi: 10.2135/cropsci2013.04.0252

Mei, D., Wang, H., Hu, Q., Li, Y., Xu, Y., and Li, Y. (2009). QTL analysis on plant height and flowering time in *Brassica napus*. *Plant Breed.* 128, 458–465. doi: 10.1111/j.1439-0523.2008.01528.x

Nordborg, M., and Tavaré, S. (2002). Linkage disequilibrium: what history has to tell us. *Trends Genet.* 18, 83–90. doi: 10.1016/S0168-9525(02)02557-X

Parkin, I. A., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., et al. (2005). Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* 171, 765–781. doi: 10.1534/genetics.105.042093

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., et al. (2014). Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLoS ONE* 9:e101673. doi: 10.1371/journal.pone.0101673

Robert, H. S., and Friml, J. (2009). Auxin and other signals on the move in plants. *Nat. Chem. Biol.* 5, 325–332. doi: 10.1038/nchembio.170

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes* 4, 137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Sang, D., Chen, D., Liu, G., Liang, Y., Huang, L., Meng, X., et al. (2014). Strigolactones regulate rice tiller angle by attenuating shoot gravitropism through inhibiting auxin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11199–11204. doi: 10.1073/pnas.1411859111

Shi, J., Li, R., Qiu, D., Jiang, C., Long, Y., Morgan, C., et al. (2009). Unraveling the complex trait of crop yield with quantitative trait loci mapping in *Brassica napus*. *Genetics* 182, 851–861. doi: 10.1534/genetics.109.101642

Snowdon, R. J., and Iniguez Luy, F. L. (2012). Potential to improve oilseed rape and canola breeding in the genomics era. *Plant Breed.* 131, 351–360. doi: 10.1111/j.1439-0523.2012.01976.x

Song, X., and Zhang, T. (2009). Quantitative trait loci controlling plant architectural traits in cotton. *Plant Sci.* 177, 317–323. doi: 10.1016/j.plantsci.2009.05.015

Stich, B., and Melchinger, A. E. (2009). Comparison of mixed-model approaches for association mapping in rapeseed, potato, sugar beet, maize, and Arabidopsis. *BMC Genomics* 10:94. doi: 10.1186/1471-2164-10-94

Teichmann, T., and Muhr, M. (2015). Shaping plant architecture. *Front. Plant Sci.* 6:233. doi: 10.3389/fpls.2015.00233

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., et al. (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* 43, 159–162. doi: 10.1038/ng.746

Turner, S. D. (2014). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *biorXiv*. doi: 10.1101/005165

Wang, N., Li, F., Chen, B., Xu, K., Yan, G., Qiao, J., et al. (2014). Genome-wide investigation of genetic changes during modern breeding of *Brassica napus*. *Theor. Appl. Genet.* 127, 1817–1829. doi: 10.1007/s00122-014-2343-6

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Wang, Y., and Li, J. (2005). The plant architecture of rice (*Oryza sativa*). *Plant Mol. Biol.* 59, 75–84. doi: 10.1007/s11103-004-4038-x

Wang, Y., and Li, J. (2008). Rice, rising. *Nat. Genet.* 40, 1273–1275. doi: 10.1038/ng1108-1273

Xu, J., Song, X., Cheng, Y., Zou, X., Zeng, L., Qiao, X., et al. (2014). Identification of QTL for branch number in oilseed rape (*Brassica napus* L.). *J. Genet. Genomics* 41, 557–559. doi: 10.1016/j.jgg.2014.06.006

Yoshihara, T., Spalding, E. P., and Iino, M. (2013). AtLAZY1 is a signaling component required for gravitropism of the *Arabidopsis thaliana* inflorescence. *Plant J.* 74, 267–279. doi: 10.1111/tpj.12118

Yu, B., Lin, Z., Li, H., Li, X., Li, J., Wang, Y., et al. (2007). TAC1, a major quantitative trait locus controlling tiller angle in rice. *Plant J.* 52, 891–898. doi: 10.1111/j.1365-313X.2007.03284.x

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., et al. (2005). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702

Ziolkowski, P. A., Kaczmarek, M., Babula, D., and Sadowski, J. (2006). Genome evolution in Arabidopsis/Brassica: conservation and divergence of ancient rearranged segments and their breakpoints. *Plant J.* 47, 63–74. doi: 10.1111/j.1365-313X.2006.02762.x

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355–360. doi: 10.1038/ng.546

Zhao, K., Aranzana, M. J., Kim, S., Lister, C., Shindo, C., Tang, C., et al. (2007). An Arabidopsis example of association mapping in structured samples. *PLoS Genet.* 3:e4. doi: 10.1371/journal.pgen.0030004

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2:467. doi: 10.1038/ncomms1467

Zhao, S. Q., Xiang, J. J., and Xue, H. W. (2012). Studies on the rice leaf INCLINATION1 (LC1), an IAA–amido synthetase, reveal the effects of auxin in leaf inclination control. *Mol. Plant* 6, 174–187. doi: 10.1093/mp/sss064

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Liu, Wang, Mei, Wang, Fu, Liu, Li and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 March 2016
doi: 10.3389/fpls.2016.00288

Intraspecific Variability of Floral Nectar Volume and Composition in Rapeseed (*Brassica napus* L. var. *oleifera*)

Michele Bertazzini and Giuseppe Forlani **

Department of Life Science and Biotechnology, University of Ferrara, Ferrara, Italy

Numerous angiosperms rely on pollinators to ensure efficient flower fertilization, offering a reward consisting of nourishing nectars produced by specialized floral cells, known as nectaries. Nectar components are believed to derive from phloem sap that is enzymatically processed and transformed within nectaries. An increasing body of evidence suggests that nectar composition, mainly amino acids, may influence pollinator attraction and fidelity. This notwithstanding, little is known about the range of natural variability in nectar content for crop species. Sugar and amino acid composition of nectar harvested from field-grown plants at the 63–65 phenological stage was determined for a set of 44 winter genotypes of rapeseed, a bee-pollinated crop. Significant differences were found for solute concentrations, and an even higher variability was evident for nectar volumes, resulting in striking differences when results were expressed on a single flower basis. The comparison of nectar and phloem sap from a subset of eight varieties pointed out qualitative and quantitative diversities with respect to both sugars and amino acids. Notably, amino acid concentration in phloem sap was up to 100 times higher than in nectar. Phloem sap showed a much more uniform composition, suggesting that nectar variability depends mainly on nectary metabolism. A better understanding of the basis of nectar production would allow an improvement of seed set efficiency, as well as hive management and honey production.

Keywords: nectar production, phloem sap, amino acid and sugar content, nectary metabolism, honeybee preference

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Bożena Denisow,
University of Life Sciences in Lublin, Poland
Vartika Mathur,
University of Delhi, India

***Correspondence:**
Giuseppe Forlani
flg@unife.it

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *24 December 2015*
Accepted: *23 February 2016*
Published: *16 March 2016*

Citation:
Bertazzini M and Forlani G (2016) Intraspecific Variability of Floral Nectar Volume and Composition in Rapeseed (Brassica napus L. var. oleifera). Front. Plant Sci. 7:288. doi: 10.3389/fpls.2016.00288

INTRODUCTION

Many plants require pollinator visitation to obtain efficient seed set. Dicotyledonous species often attract pollinators by offering them a reward of floral nectars. The nectar is a nutrient-rich aqueous solution of sugars, amino acids, organic acids, proteins, fats, vitamins, minerals, and other minor components, such as proteins with high antimicrobial activity (Nicolson and Thornburg, 2007). It is derived from the phloem sap and is produced by a group of specialized cells, the nectaries, usually present in the flower at the base of the petals (De La Barrera and Nobel, 2004). Besides, nectar may also contain secondary metabolites such as terpenes, alkaloids, flavonoids, vitamins, and oils (Truchado et al., 2008). Nectar has a significant metabolic cost for the plant, so as to often be resorbed once fertilization has occurred (Nepi and Stpiczyńska, 2008). Its composition can vary greatly depending on plant species and environmental conditions (Herrera et al., 2006), as well

as on floral sexual phases (Antoñ and Denisow, 2014), and flower position within inflorescences (Lu et al., 2015). Total sugar content ranges from a minimum of 5% to a maximum of 80%. In most cases, sucrose is the only or main component, but in some cases sucrose, glucose and fructose are present in similar amounts, conditioning pollinator choice (Lotz and Schondube, 2006). Only rarely other monosaccharides such as raffinose, galactose, and sorbitol are found. Since the phloem sap contains mostly sucrose, chemical reactions must occur to produce glucose and fructose in the nectar. Unequivocal data have been reported showing that these reactions are catalyzed by transglucosidases and transfructosidases localized in the nectaries (Heil, 2011).

Amino acids are also found in nectar but at much lower quantities (typically 0.02–4.8% organic matter), and the biological significance of their presence is still being debated. Some authors have shown that plants pollinated by butterflies contain a higher concentration of amino acids in their nectar than those pollinated by birds (Baker and Baker, 1973). It was believed that amino acid composition was constant in plants of the same species, even if grown in different environments (Baker and Baker, 1977). However, this concept has been superseded by experimental evidence showing a high variability both inter- and intra-populations, strongly influenced by nitrogen availability (Gardener and Gillman, 2001b). The quantity and quality of amino acids in the nectar may enhance insect longevity and fecundity. Radiotracer studies showed that amino acids ingested by adults are incorporated into eggs. Female butterflies prefer nectars spiced with amino acids to nectars lacking them, whereas males show no preference (Mevi-Schutz and Erhardt, 2005).

Among amino acids found in nectars, proline has a unique feature: it is capable of stimulating the labellar salt receptor cells of some insect species, which therefore seem able to recognize the taste (Hansen et al., 1998; Wacht et al., 2000). Nectar foraging insects preferentially utilize proline during the initial phases of flight (Micheu et al., 2000). The availability of an energetic substrate ready to be used and suitable for intense flight phases can represent an advantage for bees during long distance foraging. Proline represents the most abundant amino acid in the haemolymph of many insects, including honeybees (Crailsheim and Leonhard, 1997), and high amounts of proline are found in many types of nectar. In tobacco plants it can accumulate to levels of 45–60% of total amino acids (Carter et al., 2006). In addition to proline, aromatic amino acids (tyrosine, phenylalanine), serine, and amides (glutamine, asparagine) may also be present at high concentrations (Gardener and Gillman, 2001a). Increasing evidence supports the preference of bees and butterflies for sugar solutions enriched with proline. The concentration range preferred by honeybees (from 2 to 6 mM; Carter et al., 2006) is close to that found in several natural nectars (Gardener and Gillman, 2001a). Such a preference does not seem to exist in bird pollinators (Leseigneur et al., 2007). This suggests a co-evolutionary strategy for increasing pollination of plants that produce proline-rich nectar by foraging insects that perceive its presence (Biancucci et al., 2015). Our earlier studies found a strong preference of proline-rich nectar and aversion to nectar containing serine by forager honeybees (Bertazzini et al., 2010), and these preferences and aversions may influence the frequency of flower visitation by insects.

Similarly, the quality and the ratios between various types of sugars and their absolute amounts *per* flower may greatly alter the attractiveness to pollinators. In the case of honeybees, it is well-known that the profitability of a flower, defined as the ratio between the caloric cost to fly to and visit a single flower and the mean caloric gain that can be obtained while foraging, is one of the main determinants of flower choice and of dance communication of scout honeybees to hivemates (Waddington, 1982). Amino acid and sugar content may thus contribute in providing the basis for flower constancy, the phenomenon by which an individual forager actually bypasses rewarding flowers to restrict visits to a single plant species (Grüter and Ratnieks, 2011). A better knowledge of these aspects may open new perspectives in both hive management and optimization of crop yield. The occurrence of a natural variability in nectar composition among cultivars of a bee-pollinated crop could cause a different frequency of visit by pollinators, resulting in different seed set efficiency, and significantly influencing final grain harvest. Moreover, positioning hives near a field where a preferred nectar-producing crop variety is cultivated could *persuade* the bees to visit this source of nectar. Feeding on a single plant species, bees would produce a more valued (*unifloral*) honey, with a distinctive aroma and flavor.

Many crops are dependent to different degrees on honeybees for pollination. These include apples, avocados, cherries, cranberries, sunflowers, alfalfa, cucumbers, kiwi fruit, and melons (Ball, 2007). Oilseed rape (*Brassica napus* L. var. *oleifera*), despite being considered a predominantly wind-pollinated and self-compatible plant, in a number of studies showed significantly increased grain yields when bee-pollinated (Chambó et al., 2014, and the references therein). Winter rapeseed cultivation has dramatically increased in Europe following the 2003/30 Directive of the European Parliament promoting the use of biofuels to replace diesel or gasoline for transport. Almost one-third of the total cropped area has been reported to be entomophilous, and pollinators not only double the final yield, but also contribute to achieve uniform and early pod setting (Abrol, 2007). Several studies investigated sugar composition of rapeseed nectar (reviewed in Westcott and Nelson, 2001). In most cases sucrose was present at very low levels, together with high fructose concentrations (e.g., Kevan et al., 1991). The presence of a low fructose-to-glucose ratio in unifloral rapeseed honey is the cause of its high tendency to granulate, a property that forces beekeepers to harvest it as soon as it is capped. After collecting, it has to be extracted within 24 h and marketed within a few weeks. Despite these unfavorable features, rapeseed shows one of the highest melliferous potentials and represents a main forage crop for bees. Many palynologic analyses showed the presence of notable percentages of rapeseed pollen in multifloral honeys (e.g., Sabo et al., 2011). Yet, a recent calculation of the theoretical maximal honey yield revealed that this bee pasture may be considerably underutilized (Nedić et al., 2013).

The occurrence of a significant natural variability in sugar and amino acid content in nectar would allow breeding programs to increase both the attractiveness to honeybees (influencing in

turn crop yield) and honey quality. With the aim to confirm previous data on sugar content and obtain new information on amino acid composition, we harvested and analyzed nectars from a large group of rapeseed varieties. To investigate the relative role of phloematic sap and nectary metabolism in establishing the nature and quantity of sugars and amino acids in floral nectars, the phloem sap of a subset of genotypes was also analyzed.

MATERIALS AND METHODS

Plant Growth

Seeds of 44 commercial winter cultivars of *B. napus* L. var. *oleifera* Metzger (including 16 inbred lines and 28 hybrid varieties, as detailed in **Table 1**) were sown on October 3rd, 2009, in an experimental field located near Jolanda di Savoia (FE), Italy, at 44°54′42.3″ N–11°58′44.4″ E, which during the previous year had not been cultivated. Soil properties were as summarized in Table S1. A completely randomized design with four replicates was adopted (Figure S2). The field was divided into four parts. Each part enclosed 44 plots (2 × 2 m), each one consisting of seven rows, 33 cm apart. A 1.0 m edge was left all around. Seeds were sown with a mechanical planter (1001-B Precision Garden Seeder, Earthway, equipped with a 1002/05 seed plate), obtaining a density of about 50 seedlings m^{-1}. Fertilization consisted of 40 kg ha^{-1} N (urea) and 5 kg ha^{-1} P_2O_5 (superphosphate) in pre-emergence, and 80 + 40 kg ha^{-1} N (ammonium nitrate) topdressed at the flower-bud-visibility stage. Irrigation was not used, and no chemical treatment was given in order to limit weed growth. Weeds were removed manually, when required. Immediately before nectar sampling, three plots with plants showing a uniform growth were selected for each rapeseed variety (Figure S2), whereas plants in the fourth plot were not used. Meteorological data (irradiance, temperature, rainfall, relative humidity, and pressure) for the study period are reported in Figure S3.

Nectar Sampling

Flowers were sampled from plants at the 63 (30% of flowers on main raceme open) to 65 (50% flowers on main raceme open, older petals falling) phenological stage in the BBCH-scale (Lancashire et al., 1991). Nectar was extracted by centrifugation with the method of Bosi (1973). For each sample, 40 freshly-opened flowers were harvested, one flower *per* plant, taking care to discard those already visited by foraging insects. Flowers were transferred into sterile 50-mL centrifugal tubes containing 10 g of acid-washed glass beads (5 mm diameter, Sigma 18406) wrapped in a nylon mesh and covered by a piece of hydrophobic cotton to avoid sample contamination by pollen. Following centrifugation

TABLE 1 | Rapeseed varieties used in this study.

Cultivar	Source[a]		Cultivar	Source[a]		Cultivar	Source[a]	
Adam	1	Inbred	Gamin	9	Hybrid	Pelican	3	Hybrid
Alpaga	2	Inbred	Henry	10	Inbred	Pr45d01	12	Hybrid
Aragon	3	Inbred	Hercules	4	Hybrid	Pr45d03	12	Hybrid
Avenir	2	Hybrid	Hybristar	8	Hybrid	Pr46w10	12	Hybrid
Bambin	2	Hybrid	Intense	2	Hybrid	Pr46w31	12	Hybrid
Belcanto	2	Hybrid	Katabatic	11	Inbred	Shakira	10	Inbred
Beluga	2	Inbred	Kompass	1	Hybrid	Taurus	3	Hybrid
Dante	4	Inbred	Lorenz	3	Inbred	Toccata	11	Hybrid
Elan	5	Hybrid	Mendel	3	Hybrid	Totem	2	Inbred
Es Artist	6	Hybrid	Milena	5	Inbred	Vectra	4	Hybrid
Exagone	7	Hybrid	Nk Caravel	9	Hybrid	Verona	13	Inbred
Excalibur	7	Hybrid	Nk Formula	9	Hybrid	Viking	3	Inbred
Excel	7	Hybrid	Nk Petrol	9	Hybrid	Zeruca	2	Inbred
Facile	3	Hybrid	Nk Technic	9	Hybrid	Zoom	2	Inbred
Forza	8	Inbred	Palmedor	2	Hybrid			

[a] *1 Deutsche Saatveredelung Lippstadt, courtesy of Venturoli Sementi.*
2 Serasem, courtesy of Florisem Italia.
3 Norddeutsche Pflanzenzucht Lembke Semences, courtesy of F.lli Moretti Cereali and Florisem.
4 RAPS GBR Saatzucht Lundsgaard, courtesy of Carla Sementi.
5 KWS SAAT AG, courtesy of Fondazione per l'Agricoltura F.lli Navarra.
6 Euralis Semences International, courtesy of Fondazione per l'Agricoltura F.lli Navarra.
7 Dekalb—Monsanto Company, courtesy of Monsanto Italia.
8 SCA Adrien Momont et Fils, courtesy of Istituto Sementi e Tecnologie Agroalimentari.
9 Syngenta Seeds, courtesy of Società Italiana Sementi and NK Sementi Syngenta.
10 Intersaatzucht Donau GMBH & Co., courtesy of Padana Sementi Elette.
11 Mick Pickford, courtesy of Maisadour Semences Italia.
12 Pioneer Hi-Bred, courtesy of Pioneer Hi-Bred Italia.
13 Svalof Weibull Ab, courtesy of Padana Sementi Elette.
Underlined varieties were also used to harvest phloem sap.

at RT for 3 min at 1000 g, beads and cotton were removed and the nectar was harvested, measured with a micropipette (0.5–10 μL) and transferred into 0.5-mL Eppendorf vials. Samples were immediately frozen and stored at −20°C until the analysis. Nectars were harvested in the afternoon (from 1.30 to 5.30 p.m.), and harvesting was carried out during 4 different days, on April 21st, April 25th, April 27th, and April 29th, 2010. Each time, three replications were carried out for a given genotype, one *per* part. Overall, 12 nectar samples were harvested for each rapeseed variety (3 replications [part] × 4 harvest days).

Phloem Sampling

Phloem sap was harvested with a validated protocol for species belonging to the genus *Brassica* (Giavalisco et al., 2006). Samples were obtained by making small punctures with a sterile hypodermic needle into inflorescence stems of rapeseed plants. The first exuding droplet was discarded, and the subsequent exudate was collected, immediately frozen and stored at −20°C until the analysis. Sampling was carried out in triplication on April 29th for a subset of eight genotypes, which are underlined in **Table 1**.

Sugar Analysis

Glucose, fructose, and sucrose from nectar were determined enzymatically (Figure S4). To measure glucose concentration in nectars, sample aliquots (0.1 and 0.2 μL) were incubated at 37°C in a final volume of 1 mL with 0.75 mM NAD^+, 0.5 mM ATP, and 25 mU of both hexokinase and glucose-6-phosphate dehydrogenase (Sigma G3293) in 25 mM Tris-HCl buffer, pH 7.8. The increase in absorbance at 340 nm was followed in a Peltier-equipped spectrophotometer for 15 min, until it stabilized. Fructose content was then quantified by adding 600 mU of phosphoglucose isomerase (Sigma F2668) in a volume of 10 μL. The resulting increase in absorbance was measured for a further 21 min. Sucrose concentration in the same sample was lastly evaluated by adding 30 U of baker yeast invertase (Sigma I4504) in a volume of 10 μL, monitoring the absorbance for a further 24 min. Sugar content was calculated on the basis of calibration curves obtained with suitable dilutions of an artificial nectar composed of 7.5% (w/v, 416 mM) of both glucose and fructose, and 1% (w/v, 29.2 mM) sucrose. In the case of phloem sap, 0.1 and 0.2 μL were analyzed for glucose and fructose, whereas 0.01 and 0.02 μL were used to measure sucrose content, and concentrations were estimated from calibration curves obtained with an artificial phloem sap composed of 0.5% (w/v, 27.8 mM) of both glucose and fructose and 20% (w/v, 584 mM) sucrose.

Amino Acid Analysis

Total amino acid content was determined by its reaction with *o*-phthaldialdehyde (*o*PDA) as described previously (Jones et al., 2002), with minor modifications. Sample aliquots (1 and 2 μL) were water-diluted to 50 μL, and the resulting samples were mixed with the same volume of *o*PDA solution (0.5 M in 0.5 M sodium borate buffer, pH 10.0, containing 0.5 M β-mercaptoethanol, and 10% [v/v] methanol). After exactly 60 s, the increase in absorbance was measured at 340 nm using a UV-transparent cuvette with 10 mm optical path (UVette, Eppendorf) and a universal adapter. Amino acid content was extrapolated from a calibration curve obtained with a solution of all the 20 proteinogenic compounds (each at 1 mM but asparagine, glutamic and aspartic acid at 2 mM, and glutamine at 5 mM; Figure S5).

Single amino acids were quantified by RP-HPLC following derivatization with *o*PDA, as described (Forlani et al., 2014). Peaks were integrated by area, with variation coefficients ranging from 0.8 to 3.2%. Since *o*PDA does not react with proline, its concentration was measured either by the acid ninhydrin method (Williams and Frank, 1975), or by RP-HPLC following derivatization with 4-dimethyl-aminoazobenzene-4′-sulfonyl chloride (DABS-Cl). In the latter case, 10-μL aliquots of suitable sample dilutions were mixed with the same volumes of a 0.2 M sodium bicarbonate buffer, pH 10.0, and a 2 mg mL^{-1} solution of DABS-Cl in acetone. After 30 min incubation at 70°C, 20 μL of derivatized samples were injected into a 4.6 × 250 mm Zorbax ODS column (Rockland Technologies, Newport, DE) equilibrated with 65% solvent A (17 mM potassium phosphate buffer, pH 6.5, containing 2% [v/v] dimethylformamide), and 35% solvent B (80% methanol containing 4% [v/v] dimethylformamide). Elution proceeded at a flow rate of 60 mL h^{-1} using a computer-controlled (Data System 450; Kontron, Munchen, Germany) complex gradient from 35 to 95% solvent B, monitoring the eluate at 436 nm. Two technical replications were performed for each sample. For all rapeseed varieties, single amino acid content was determined in a sample obtained by combining the same volume of the 12 existing samples. For a subset of eight genotypes, for which larger nectar volumes were available, four samples were analyzed, each one consisting of a mixture of all three nectar samples harvested in the same day. Phloem saps were analyzed individually.

Statistical Analysis

For ANOVA and Tukey's *post-hoc* HSD analysis, the *Statistica* software package (version 7.1, StatSoft, Tulsa, OK, U.S.A.) was used. Correlation analysis and descriptive statistics were computed with the *Prism 6* software (version 6.03, GraphPad Software, Inc., San Diego, CA, U.S.A.).

RESULTS

A Great Variability is Evident among Rapeseed Genotypes Regarding Floral Nectar Volume

Nectar was harvested from plants of 44 rapeseed cultivars, comprising both hybrids, and inbred lines. Preliminary attempts showed that in all cases insect-visited flowers contain negligible residual amounts of nectar (<1μL in 40 flowers). Therefore, only freshly-opened, unvisited flowers were collected, whose opening diameter was smaller and easily distinguishable from those already visited by honeybees. Moreover, harvesting was carried out in the early to mid-afternoon, when the relative humidity was lower (Figure S3), to exclude contamination by dew and to allow the attainment of steady-state nectar production,

which is dependent on the establishment of full photosynthetic rate and phloem loading. The results showed a highly significant ($P = 0.000000$) difference among cultivars with respect to nectar volume, which ranged from 20 to 750 nL flower^{-1} (**Table 2**).

A lower but statistically relevant difference was also found with respect to the day of harvest ($P = 0.003603$). However, the interaction was not significant ($P = 0.067888$), suggesting that the variation of nectar production in different days was shared by all rapeseed genotypes, and that the relative ratio was not altered. The assessment of confidence intervals for mean nectar volumes in the 4 days of harvest (**Figure 1A**) pointed out maximal amounts on April 25th, in connection with increased values of relative humidity (Figure S3) due to a moderate rain that had occurred on April 23rd. Interestingly, when data were plotted as a function of the genetic background, hybrids showed a highly significant difference ($P = 0.000012$) from inbred lines. Confidence intervals for mean volumes (**Figure 1B**) showed that the latter produce about 50% more nectar than the former. In this case also, a minor yet significant difference ($P = 0.028708$) was evident with respect to the day of harvest, but again the interaction was not significant ($P = 0.411980$).

Nectars from Rapeseed Genotypes Show Absolute Concentrations of Sugars that are Notably Different, but the Percentage Content of Glucose, Fructose, and Sucrose is Similar

The concentrations of glucose, fructose and sucrose in the harvested samples were measured on the basis of the calibration curves reported in Figure S4. Monosaccharides were nearly equally abundant, with concentrations ranging from 200 to 750 mM, whereas only minor levels of the disaccharides were detected, in the 20–60 mM range (**Table 2**). Genotypes differed significantly regarding the content of all three sugars, being $P = 0.000019$, $P = 0.000001$, and $P = 0.000000$ for glucose, fructose and sucrose, respectively. On the contrary, inbred lines and hybrids did not differ from each other with respect to the three sugars ($P = 0.409379$, $P = 0.167797$, and $P = 0.647939$), and the overall content ($P = 0.321408$). When sugar concentrations were related to the corresponding nectar volume, no significant relationship was evident for all the compounds (**Figures 2A–C**). This suggests that the differences found with respect to nectar volume do not depend only on a variable water content deriving from either sample contamination by dew (or other floral fluids), or on a variable dilution of a uniform exudate secreted by nectaries. On the contrary, a highly significant correlation was found between glucose and fructose content (**Figure 2D**), whereas sucrose concentration was related to neither of the monosaccharides (**Figures 2E,F**).

Despite the interesting differences found with respect to absolute sugar concentrations (**Figure 3B**), if glucose, fructose and sucrose levels were expressed as percentage values, a more uniform content was evident, even though genotypes still differed statistically from each other (data not shown). Percentage content of glucose ranged from 48 to 62% total sugars, that of fructose from 33 to 48% and that of sucrose from 2 to 13% (**Figure 3A**), suggesting that the relative concentrations of these compounds are determined by mechanisms that are maintained in all genotypes. However, if the absolute content for each flower is considered, i.e. if for a given genotype the different levels of the three sugars are multiplied by the corresponding nectar volume, the variability among cultivars is even more pronounced, with glucose and fructose content ranging from 4 to 250–270 nmol flower^{-1} (**Figure 3C**). This implies that the reward for an insect foraging on a single flower differs dramatically among genotypes due to the absolute content of glucose, fructose and sucrose. Based on the data herein obtained, the caloric reward *per* flower would vary over a 50-fold range, from 0.03 joule in the case of the hybrid Pr46w10 to 1.50 joule for the inbred Shakira.

Amino Acid Content in Rapeseed Nectars Shows Different Absolute Concentrations and Relatively Uniform Percentage Values

Total amino acid concentration was much lower than sugar levels, ranging from 1 to 9 mM. A remarkably high variability was found within cultivars (**Table 2**). With respect to this trait, rapeseed genotypes were significantly different ($P < 0.0001$). On the contrary, inbred lines and hybrids did not show a different amino acid content ($P = 0.6110$ in the unpaired *t*-test with Welch's correction). If amino acid concentrations were related to the corresponding nectar volumes, no significant relationship was found (**Figure 4A**), further strengthening the possibility that the variability pointed out with respect to the amount of nectar *per* flower reflects a true difference and does not depend on a different water content. Non-significant correlations were found also between amino acid levels and glucose (not shown), fructose (**Figure 4B**), or sucrose (**Figure 4C**) concentration.

Because of the low concentrations and the limited availability of nectar, for most cultivars the analysis of individual amino acid content was carried out on a sample obtained by combining the same volume of all the 12 existing specimens. Results are shown in **Figure 5**. A striking variability was found concerning absolute concentrations (**Figure 5A**), reflecting the large variations in total amino acid content (**Table 2**). However, if data were plotted as percentage values, a much more uniform picture was obtained (**Figure 5B**). Similarly to the results found for sugar content, these data seem suggestive of shared mechanisms controlling the reciprocal ratios of free amino acids. In all cases glutamine was the predominant amino acid, accounting for about one third of total content. High percent values were also found for histidine, glutamate, asparagine, and alanine. Proline levels corresponded to around 5% of total amino acid content.

For a subset of eight genotypes, the availability of larger nectar volumes allowed a suitably replicated analysis of individual amino acid content. Tabular numerical data are reported in Table S6, and results are summarized in **Figure 6**. Even within this small set of cultivars, absolute concentrations varied greatly (**Figure 6A**), and showed a significant variability ($P = 0.0025$). On the contrary, percentage data (**Figure 6B**) did not differ significantly among genotypes ($P > 0.9999$).

TABLE 2 | Volume, sugar, and amino acid content of nectars from rapeseed varieties.

Cultivar	Volume [*] (μL flower^{-1})	Glucose [§] (mM)	Fructose [§] (mM)	Sucrose [§] (mM)	Amino acids [#] (mM)
Adam	$0.51 \pm 0.12^{i-n}$	$415 \pm 26^{b-g}$	$338 \pm 24^{b-j}$	$23 \pm 1^{a-c}$	$3.79 \pm 0.54^{a-e}$
Alpaga	$0.12 \pm 0.02^{a-g}$	796 ± 56^{j}	679 ± 40^{l}	$29 \pm 2^{a,g}$	$4.76 \pm 1.83^{a-e}$
Aragon	$0.10 \pm 0.02^{a-f}$	$441 \pm 44^{b-i}$	$338 \pm 31^{b-j}$	$46 \pm 3^{j-o}$	$7.02 \pm 1.60^{a-e}$
Avenir	$0.14 \pm 0.03^{a-h}$	$454 \pm 22^{c-i}$	$365 \pm 19^{d-j}$	$25 \pm 1^{a-d}$	$4.35 \pm 1.48^{a-e}$
Bambin	$0.15 \pm 0.04^{a-h}$	$409 \pm 8^{b-g}$	$345 \pm 8^{c-j}$	$25 \pm 1^{a-d}$	$7.38 \pm 2.01^{c-e}$
Belcanto	$0.16 \pm 0.02^{a-h}$	578 ± 56^{i}	$428 \pm 39^{j,k}$	$42 \pm 4^{h-n}$	$3.55 \pm 0.85^{a-e}$
Beluga	$0.15 \pm 0.02^{a-h}$	$481 \pm 12^{e-i}$	$410 \pm 8^{h-k}$	$23 \pm 1^{a-c}$	$4.62 \pm 1.36^{a-e}$
Dante	$0.24 \pm 0.06^{a-i}$	$334 \pm 10^{a-d}$	$280 \pm 11^{a-f}$	18 ± 1^{a}	$4.53 \pm 1.10^{a-e}$
Elan	$0.12 \pm 0.03^{a-g}$	$373 \pm 33^{b-f}$	$265 \pm 23^{a-e}$	$40 \pm 3^{g-m}$	$1.32 \pm 0.17^{a,b}$
Es Artist	$0.40 \pm 0.08^{f-m}$	$445 \pm 35^{b-i}$	$343 \pm 25^{b-j}$	$48 \pm 3^{k-o}$	$1.96 \pm 0.38^{a-c}$
Exagone	$0.39 \pm 0.06^{d-m}$	$527 \pm 35^{g-i}$	$327 \pm 27^{b-j}$	$46 \pm 2^{j-o}$	$3.55 \pm 0.58^{a-e}$
Excalibur	$0.31 \pm 0.06^{a-j}$	$529 \pm 8^{g-i}$	$365 \pm 10^{d-j}$	$54 \pm 4^{n,o}$	$2.84 \pm 0.48^{a-c}$
Excel	$0.27 \pm 0.04^{a-i}$	$563 \pm 17^{h,i}$	$337 \pm 13^{b-j}$	$40 \pm 1^{f-m}$	$1.34 \pm 0.15^{a,b}$
Facile	$0.20 \pm 0.04^{a-i}$	$530 \pm 9^{g-i}$	481 ± 5^{k}	$34 \pm 3^{c-j}$	$7.54 \pm 1.19^{c-e}$
Forza	$0.68 \pm 0.12^{m,n}$	$454 \pm 13^{c-i}$	$404 \pm 11^{h-k}$	$19 \pm 2^{a,b}$	$2.90 \pm 0.98^{a-d}$
Gamin	$0.07 \pm 0.02^{a-e}$	$401 \pm 22^{b-g}$	$266 \pm 18^{a-e}$	$43 \pm 1^{i-n}$	$7.17 \pm 0.66^{b-e}$
Henry	$0.66 \pm 0.14^{l-n}$	$412 \pm 29^{b-g}$	$334 \pm 19^{b-j}$	$22 \pm 3^{a-c}$	$3.60 \pm 0.84^{a-e}$
Hercules	$0.11 \pm 0.03^{a-g}$	$391 \pm 13^{b-g}$	$315 \pm 11^{b-j}$	$32 \pm 2^{b-i}$	$4.58 \pm 1.28^{a-e}$
Hybristar	$0.39 \pm 0.06^{e-m}$	$481 \pm 13^{e-i}$	$409 \pm 15^{h-k}$	$25 \pm 1^{a-d}$	0.98 ± 0.18^{a}
Intense	$0.06 \pm 0.02^{a-c}$	$384 \pm 14^{b-f}$	$328 \pm 12^{b-j}$	$25 \pm 1^{a-d}$	$6.08 \pm 1.69^{a-e}$
Katabatic	$0.06 \pm 0.02^{a-c}$	$374 \pm 13^{b-f}$	$319 \pm 7^{b-j}$	$26 \pm 2^{a-e}$	$2.48 \pm 0.78^{a-c}$
Kompass	$0.21 \pm 0.03^{a-i}$	$441 \pm 10^{b-i}$	$354 \pm 9^{d-j}$	$23 \pm 2^{a-c}$	$5.34 \pm 1.24^{a-e}$
Lorenz	$0.64 \pm 0.09^{k-n}$	$309 \pm 11^{a,b}$	$231 \pm 2^{a,b}$	$38 \pm 1^{e-l}$	$2.61 \pm 0.60^{a-c}$
Mendel	$0.34 \pm 0.05^{a-k}$	$400 \pm 10^{b-g}$	$299 \pm 7^{b-h}$	$29 \pm 2^{a-h}$	$5.46 \pm 1.71^{a-e}$
Milena	$0.43 \pm 0.08^{g-m}$	$447 \pm 54^{b-i}$	$320 \pm 36^{b-j}$	$39 \pm 3^{f-m}$	$6.17 \pm 1.16^{a-e}$
Nk Caravel	$0.35 \pm 0.05^{b-l}$	$512 \pm 11^{f-i}$	$388 \pm 11^{f-k}$	$39 \pm 2^{f-m}$	$2.83 \pm 0.60^{a-c}$
Nk Formula	$0.09 \pm 0.01^{a-f}$	$472 \pm 20^{d-i}$	$369 \pm 18^{e-k}$	56 ± 3^{o}	$8.84 \pm 3.00^{d-e}$
Nk Petrol	$0.59 \pm 0.07^{j-n}$	$428 \pm 28^{b-h}$	$329 \pm 21^{b-j}$	$49 \pm 0^{l-o}$	$6.15 \pm 1.42^{a-e}$
Nk Technic	$0.20 \pm 0.04^{a-i}$	$442 \pm 5^{b-i}$	$330 \pm 6^{b-j}$	$39 \pm 1^{e-m}$	$1.92 \pm 0.21^{a-c}$
Palmedor	$0.20 \pm 0.03^{a-i}$	$380 \pm 1^{b-f}$	$307 \pm 4^{b-i}$	$28 \pm 1^{a-g}$	$4.09 \pm 0.99^{a-e}$
Pelican	$0.37 \pm 0.06^{c-m}$	$483 \pm 8^{e-i}$	$373 \pm 10^{e-k}$	$36 \pm 3^{d-k}$	$2.94 \pm 0.86^{a-d}$
Pr45d01	$0.07 \pm 0.01^{a-d}$	$328 \pm 8^{a-c}$	$255 \pm 10^{a-d}$	$35 \pm 1^{c-j}$	$3.67 \pm 1.23^{a-e}$
Pr45d03	$0.04 \pm 0.01^{a,b}$	$345 \pm 11^{a-e}$	$236 \pm 4^{a-c}$	$27 \pm 2^{a-f}$	$4.00 \pm 1.88^{a-e}$
Pr46w10	0.02 ± 0.01^{a}	218 ± 20^{a}	183 ± 19^{a}	$34 \pm 0^{c-j}$	Not determined
Pr46w31	$0.11 \pm 0.03^{a-f}$	$412 \pm 29^{b-g}$	$313 \pm 23^{b-i}$	$53 \pm 6^{n,o}$	$2.00 \pm 0.43^{a-c}$
Shakira	0.75 ± 0.10^{n}	$346 \pm 7^{a-e}$	$298 \pm 6^{b-h}$	$26 \pm 1^{a-e}$	$1.21 \pm 0.18^{a,b}$
Taurus	$0.45 \pm 0.07^{h-n}$	$456 \pm 24^{c-i}$	$400 \pm 26^{g-k}$	$25 \pm 0^{a-d}$	$3.20 \pm 1.08^{a-d}$
Toccata	$0.17 \pm 0.04^{a-h}$	$368 \pm 8^{b-f}$	$290 \pm 6^{a-g}$	$46 \pm 1^{j-o}$	$3.50 \pm 0.92^{a-e}$
Totem	$0.03 \pm 0.01^{a,b}$	$432 \pm 16^{b-h}$	$383 \pm 13^{f-k}$	$25 \pm 2^{a-d}$	$4.58 \pm 1.26^{a-e}$
Vectra	$0.12 \pm 0.02^{a-g}$	$425 \pm 31^{b-h}$	$309 \pm 19^{b-i}$	$30 \pm 2^{a-h}$	$3.25 \pm 0.86^{a-e}$
Verona	$0.20 \pm 0.06^{a-i}$	$416 \pm 6^{b-g}$	$331 \pm 12^{b-j}$	$35 \pm 2^{c-j}$	$1.21 \pm 0.13^{a,b}$
Viking	$0.23 \pm 0.04^{a-i}$	$373 \pm 8^{b-f}$	$283 \pm 14^{a-f}$	$40 \pm 2^{f-m}$	$2.80 \pm 0.26^{a-c}$
Zeruca	$0.06 \pm 0.02^{a-c}$	747 ± 67^{j}	607 ± 66^{l}	$51 \pm 3^{m-o}$	9.15 ± 2.45^{e}
Zoom	$0.30 \pm 0.10^{a-j}$	$445 \pm 10^{b-i}$	$419 \pm 12^{i-k}$	$28 \pm 4^{a-g}$	$1.63 \pm 0.23^{a-c}$

[*] *Samples composed of 40 flowers were collected, and nectar was harvested by centrifugation and quantified as described in Section Nectar Sampling. Results are mean ± SE over 12 replicates.*

[§] *Sugar content was determined enzymatically, as described in Section Sugar Analysis.*

[#] *Total amino acid content was quantified by reaction with oPDA, as described in Section Amino Acid Analysis.*

For a given parameter (nectar volume, glucose, fructose, sucrose, and total amino acid content), data were subjected to one-way ANOVA with post-hoc comparison using Tukey's honest significant difference test. In each column, means with the same letter are not significantly different ($P > 0.05$) from each other.

Sugar and Amino Acid Content of Phloem Sap from a Subset of Eight Rapeseed Cultivars Shows a Remarkable Uniformity, and both Absolute and Relative Values are Completely Different from those Found in Nectar

Since floral nectar derives from phloem sap, the composition of the latter was investigated in the case of the small group of genotypes for which a complete analysis of sugar and amino acid content of nectar had been obtained. Data on sucrose, glucose, fructose, and total amino acid concentrations in phloem sap are presented in **Table 3**. As expected, sucrose was the main saccharide, accounting for more than 80% of total sugars, whereas monosaccharides were present at a tenfold lower concentration. Interestingly, free amino acid content was similarly high, ranging from 180 to 250 mM.

When sugar and amino acid content in phloem sap of a given cultivar was related to the corresponding concentration in nectar, significant differences were evident (**Figure 7**). If absolute sugar concentrations were considered, data are suggestive of a major hydrolysis of sucrose in its two components by nectaries, the overall content being comparable (700–1000 mM, as expressed in monosaccharide equivalents). On the contrary, high amino acid concentrations in sap were strongly reduced in nectar. Moreover, unlike nectar, both sugar and amino acid concentrations in sap samples were remarkably similar in all tested varieties, the differences being statistically non-significant ($P = 0.3394$ for sugars in one-way ANOVA for matched measures, and $P = 0.1494$ including also amino acid content).

The last result was further confirmed when single amino acid levels were determined in phloem sap (Table S7). Contrary to previous results for nectar (**Figure 6**), and despite a much lower intra-genotype variability, differences among rapeseed varieties were not significant either when expressing single amino acid content as absolute concentrations (**Figure 8A**, $P = 0.1932$),

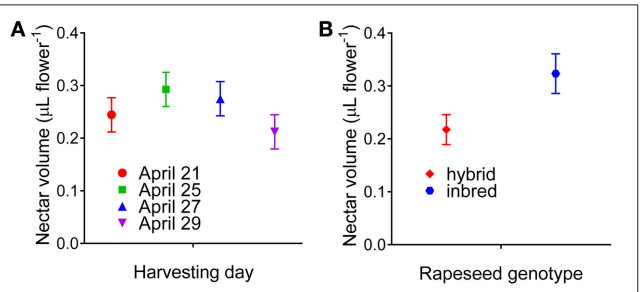


FIGURE 1 | Confidence intervals for mean nectar production. Factorial ANOVA pointed out the occurrence of significant differences in nectar production by rapeseed cultivars also as a function of the day of harvest **(A)** and their genetic nature [hybrids vs. inbred lines **(B)**]. Reported intervals are at the 95% confidence level.

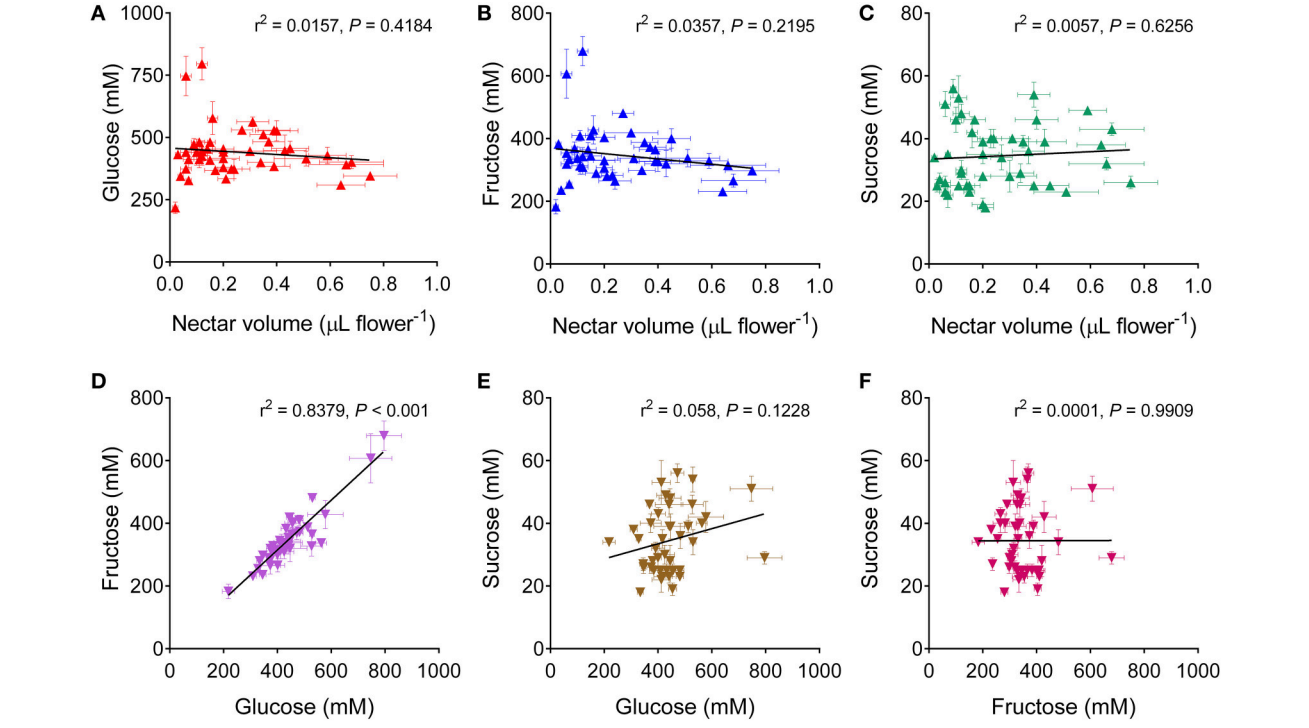


FIGURE 2 | Correlation analysis of sugar content in rapeseed nectar. Glucose, fructose and sucrose concentrations in the nectar of a given cultivar were plotted against each other **(D–F)**, or against the corresponding nectar volume **(A–C)**. In all cases a two-tailed Pearson correlation test showed no significant relationship, with the only exception of glucose and fructose content, which were found to be highly correlated **(D)**.

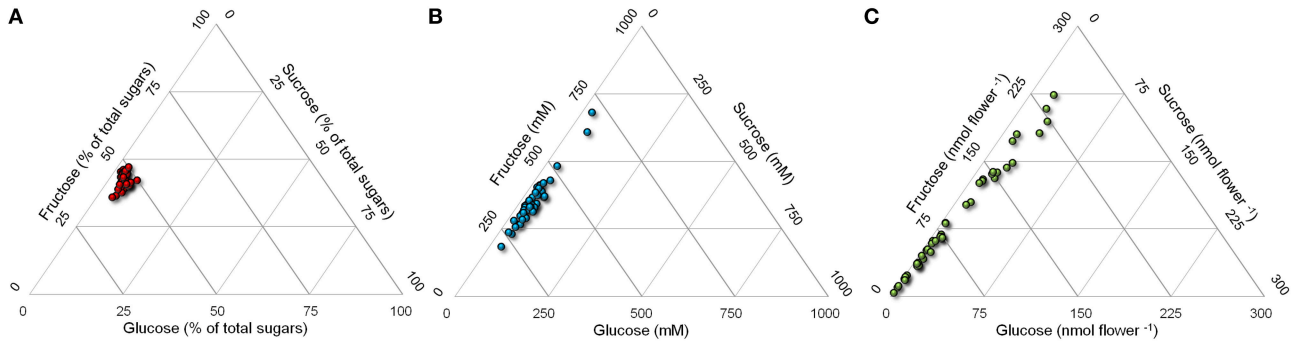


FIGURE 3 | Ternary plot of sugar content in rapeseed nectar. Glucose, fructose, and sucrose levels in nectars from 44 rapeseed cultivars were plotted in triangular graphs depicting their relative ratios. Data are expressed as percentage content **(A)**, millimolar concentration **(B)**, or absolute amount *per* flower **(C)**.

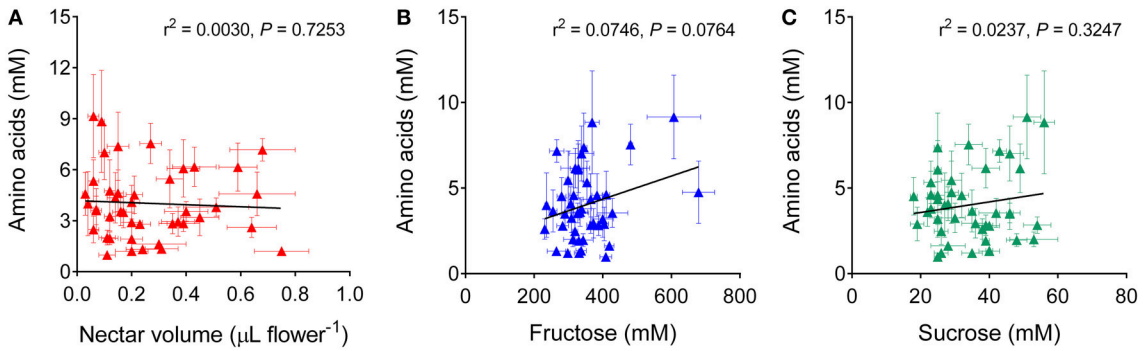


FIGURE 4 | Correlation analysis of amino acid content in rapeseed nectar. Total amino acid concentration in the nectar of a given cultivar was plotted against the corresponding nectar volume **(A)**, or against fructose **(B)**, or sucrose **(C)** content. In all cases a two-tailed Pearson correlation test showed no significant relationship.

or when plotting data as percentage values (**Figure 8B**, $P = 0.9969$). Considering variations between nectar and phloem sap, absolute differences were not very informative, nectar content being strongly reduced in all cases. If expressed as percentage differences, data showed that some amino acids were proportionally enriched in nectars, whereas others were reduced (**Figure 9**). Among the latter the main variation concerned glutamine, whose contribution to total amino acids halved. Proline content increased slightly, similarly to several other amino acids, among which aspartic and glutamic acid, asparagine, and γ-aminobutyric acid.

DISCUSSION

This study aimed at investigating the occurrence of intraspecific variability among cultivated rapeseed winter genotypes with respect to sugar and amino acid content in nectar. Sugar concentration and quality can significantly influence both foraging insect preference and, in the case of bees, the properties of the resulting honey. For instance, a high glucose-to-fructose ratio can cause increased tendency of honey to granulate even in the hives, forcing beekeepers to adopt special process of harvesting. Moreover, honeybee preference may positively affect also seed set efficiency, possibly leading to significant increase of crop yield (Abrol and Shankar, 2012). Several previous reports on rapeseed nectar showed no or very low sucrose concentrations, and a relatively low fructose-to-glucose ratio, in the range 0.80–0.95 (Westcott and Nelson, 2001). These data were generally obtained on a small number of rapeseed cultivars (e.g., Mohr and Jay, 1990). Only a very few studies to date have investigated nectar properties in a wide array of rapeseed genotypes (Kevan et al., 1991; Pierre et al., 1999). In those surveys only limited characterization was performed. In the former study, only glucose and fructose concentrations in nectars of 25 cultivars were considered (Kevan et al., 1991). In the latter case, the variability in nectar secretion among 71 genotypes of winter oilseed rapes was tested for floral nectar volume, sugar composition, and concentration (Pierre et al., 1999). In both cases, neither amino acid level, nor the relationship between sugar and amino acid content of nectar and phloem sap was determined.

The present results highlighted the occurrence of a remarkable variability among rapeseed genotypes concerning nectar volume, with mean values ranging from 0.02 to 0.75 µL flower^{-1}. A lower, yet significant eight-fold variation had also been reported in one of the aforementioned studies (Pierre et al., 1999). In that case, however, much higher absolute values had been

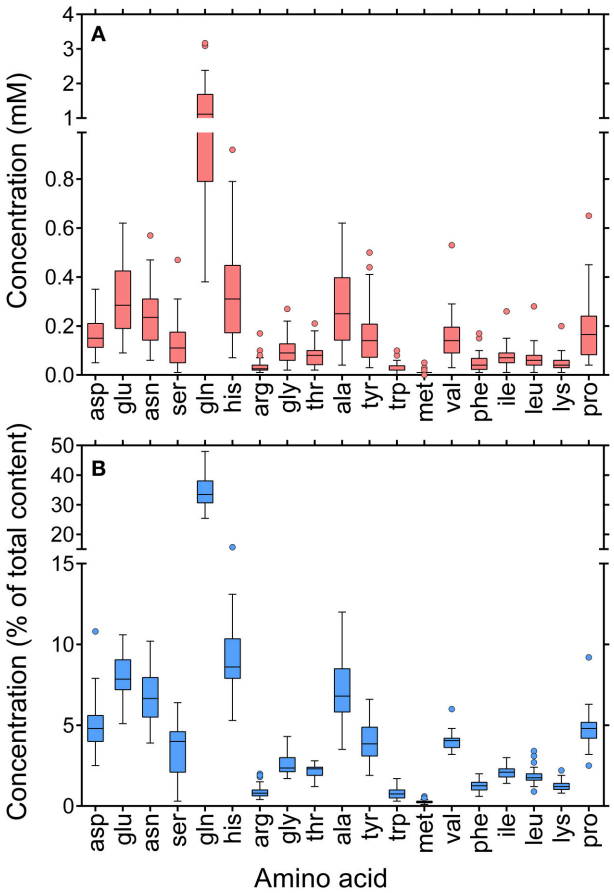


FIGURE 5 | Whisker-and-box plots of amino acid content in rapeseed nectar. Single amino acid concentrations were determined by RP-HPLC analysis of a sample obtained by combining the same volume of all the 12 existing specimens for a given genotype. Because of the limited amounts of nectar available, only 36 of 44 cultivars were analyzed. Results were expressed as either absolute concentrations **(A)**, or percent values of total amino acid content **(B)**.

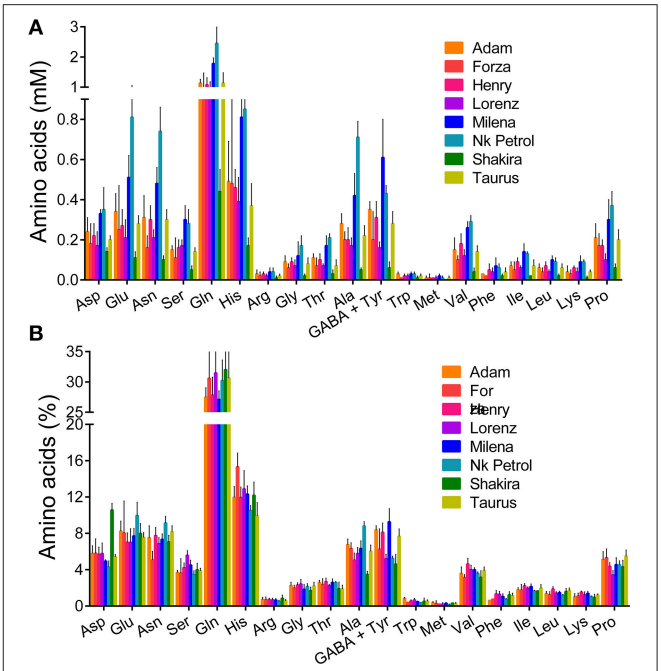


FIGURE 6 | Single amino acid content in a subset of eight rapeseed cultivars. For those genotypes for which larger amounts of nectar were available, single amino acid concentrations were determined by RP-HPLC analysis of four samples obtained by combining the three specimens harvested in the same day. Results were expressed as either absolute concentrations **(A)**, or percent values of total amino acid content **(B)**.

found, from 0.7 to 5.9 μL flower^{-1}. Such a discrepancy may depend on the adoption of a different protocol for nectar harvesting (centrifugation vs. collection with micropipettes) or, more probably, on different field conditions. In fact, in that as well as in most other previous works (e.g., Mesquida et al., 1988; Nedić et al., 2013) nectars were sampled from flowers bagged 24 h prior to sampling in order to prevent nectar uptake by foraging insects. On the contrary in the present study freshly-opened, unvisited flowers were harvested from plants that were continuously visited by bees and other foraging insects, which is a much more natural condition. It is quite likely that in completely unvisited plants nectar could accumulate with time, reaching higher volumes *per* flower.

An increasing amount of data emphasized that nectar volume and composition may be significantly influenced by a number of factors, including relative humidity, time of day, and soil composition. For instance, the flowers of some cultivars of both *B. napus* and *B. campestris* produced more nectar - with a lower sugar content- in the morning, and correlations

TABLE 3 | Sugar and amino acid content of rapeseed phloem sap.

Cultivar	Sucrose (mM)	Glucose (mM)	Fructose (mM)	Amino acids (mM)
Adam	442 ± 9^a	46.7 ± 15.3^a	38.8 ± 11.6^a	230 ± 19^a
Forza	420 ± 31^a	26.9 ± 5.3^a	22.4 ± 4.7^a	199 ± 15^a
Henry	493 ± 74^a	30.7 ± 5.4^a	25.5 ± 4.7^a	226 ± 20^a
Lorenz	440 ± 25^a	35.9 ± 3.8^a	31.3 ± 4.0^a	195 ± 16^a
Milena	476 ± 32^a	30.4 ± 2.8^a	24.9 ± 2.3^a	224 ± 16^a
Nk Petrol	446 ± 19^a	27.7 ± 2.1^a	24.3 ± 1.4^a	219 ± 28^a
Shakira	494 ± 41^a	47.5 ± 11.0^a	39.7 ± 8.9^a	249 ± 26^a
Taurus	397 ± 27^a	31.2 ± 4.8^a	27.4 ± 4.4^a	187 ± 11^a

For a given parameter (sucrose, glucose, fructose, and total amino acid content), data were subjected to one-way ANOVA with post-hoc comparison using Tukey's honest significant difference test. In each column, means with the same letter are not significantly different ($P > 0.05$) from each other.

were found between the amounts and concentrations of nectar produced and temperature or relative humidity (Mohr and Jay, 1990). In the present study, special attention was paid to standardize nectar harvesting conditions, as to obtain reproducible results and allow a proper comparison among rapeseed genotypes. Nectar collection was not carried out in the morning, when the relative humidity is higher and samples may be contamined by dew, but in the early to mid-afternoon, when the establishment of full photosynthetic rate and phloem loading should lead to the attainment of steady-state nectar production. Moreover, harvesting was performed

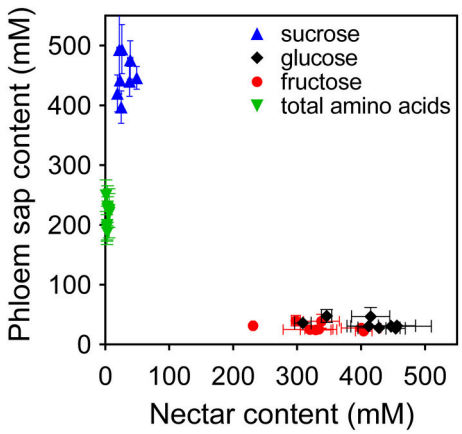


FIGURE 7 | Comparison of sugar and amino acid content in rapeseed nectar and phloem sap. The concentrations of glucose, fructose, sucrose, and total amino acid found in nectar were plotted against those found in phloem sap. Data, expressed as mean values ± SE, refer to the subset of eight genotypes for which sap had been analyzed.

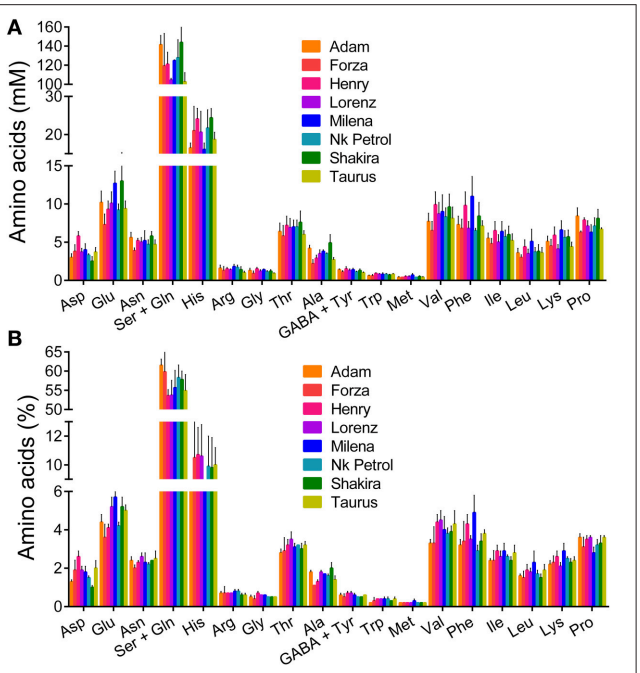


FIGURE 8 | Single amino acid content in phloem sap. For a subset of eight genotypes, single amino acid concentrations in phloem sap were determined by RP-HPLC analysis. Results are means ± SE over three biological replications, and were expressed as either absolute concentrations **(A)**, or percent values of total amino acid content **(B)**.

in sunny days, with similar temperature and humidity (Figure S3). This notwithstanding, significant variations were found between nectar volumes harvested in different days (**Figure 1A**). Consistently with literature data (Mohr and Jay, 1990), the highest values were obtained following a moderate rain that occurred after the first harvesting day, despite the fact that the second harvest had been postponed for this reason. However, lowest and highest mean values differed by <40%, whereas in other studies up to three-fold variations were reported (Pierre et al., 1999). Moreover, the results of a factorial analysis of variance showed that the increase was not dissimilar in all genotypes, strengthening the conclusion that the differences found reflect true differences in nectar production.

Besides environmental conditions, nectar content may also vary depending on floral sexual phases (Antoñ and Denisow, 2014) and flower position within inflorescences (Lu et al., 2015). However, these aspects were not considered, the aim of this study being to compare nectar from different genotypes. The use of highly standardized and uniform harvesting conditions should have minimized the effect of these variability factors. Interestingly, besides the highly significant difference among genotypes, nectar volumes from rapeseed inbred lines were found to differ significantly from those produced by hybrids, with the mean value of the latter about 50% lower than the former (**Figure 1B**). Such a difference might depend at least in part on the presence of a higher number of flowers *per* plant in hybrids, the number of pods being one of the eight yield-correlated traits showing significant mid-parent heterosis found in rapeseed (Shi et al., 2011). To the best of our knowledge, no other information is available to date regarding this point, since in the previous survey in which a significant number of rapeseed genotypes were considered, only one hybrid was included among 71 rape lines. However, in that case a significant difference with respect to nectar volume was found among three types of conventional fertile oilseed rape lines and varieties differing in seed quality, i.e. containing low (0) or high (+) levels of erucic acid and glucosinolates (00, 0+, and ++, respectively; Pierre et al., 1999). Whatever the reason for such a difference, this could result in a lower preference of honeybees for hybrid varieties, the caloric reward *per* flower being significantly reduced. If so, a lower impollination rate could partially counteract the outbreeding enhancement.

With respect to sugar content, the results herein described are on the whole consistent with those of previous reports, in all cases glucose concentration being higher than that of fructose, and sucrose concentration much lower than both of the two hexoses. However, significant levels of sucrose were found in all samples and, if expressed as percent of total sugar content, they were remarkably higher than previously reported, ranging from 3.7 to 14.3%, with a mean value of 8.2%. As a basis for comparison, Pierre et al. (1999) reported a mean sucrose content ranging from 0.3 to 0.6%. This discrepancy may depend on the different analytical techniques used (enzymatic assays vs. HPLC quantitation). However, it cannot be excluded that also in this case different results may derive from different conditions under which nectar sampling was carried out. In flowers bagged for 24 h before harvesting, nectar might be processed for a longer time by nectaries, allowing an almost complete hydrolysis of sucrose. On the contrary, in plants continuously visited by foraging bees nectar could be produced at higher speed to replace the amounts taken by insects, without the time to complete the invertase reaction. Further work will be required to discriminate between these hypotheses. Concerning

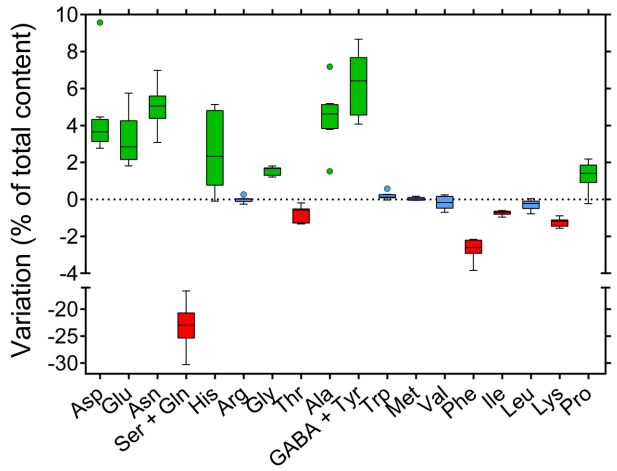


FIGURE 9 | Variation of single amino acid content between nectar and phloem sap. The levels of free amino acids found in nectar were compared with those found in phloem sap. For each compound data were expressed as the variation of its percent contribution to total amino acids. Whisker-and-box plots obtained for the subset of eight genotypes whose phloem sap had been analyzed are presented.

absolute sugar concentrations, highly significant differences were found among rapeseed genotypes, ranging from 0.468 to 1.532 M hexose equivalents. In a previous study no clear genotypic effect on sugar levels could be demonstrated. However, in that case the analysis was carried out to verify divergence among varieties differing in seed quality, and genotypes of each type were not equally represented on every harvesting date (Pierre et al., 1999). Therefore, this is the first report clearly showing the occurrence of a high variability in rapeseed regarding this trait, as shown to occur in other Brassicaceous species (Denisow et al., 2015).

This is also the first report describing free amino acid content in rapeseed nectar. With respect to both total amino acid content and single amino acid composition, the genotypes analyzed showed a high degree of diversity. Consistently with literature data on other species (e.g., Gardener and Gillman, 2001a; Carter et al., 2006), the overall absolute concentrations ranged from 1 to 10 mM. Glutamine was the most abundant, accounting for about one third of total content. Relatively high levels were also found for histidine, glutamate, asparagine, and alanine. Such a composition is quite different from the ideal ratio of essential amino acids that are required for the normal growth and development of bees (De Groot, 1953). Moreover, the content of the only amino acid that most insects have the ability to taste, proline (Wacht et al., 2000; Gardener and Gillman, 2002), was quite low in all genotypes, and below the levels (1–5 mM) that were found in nectars from other plant species (Carter et al., 2006; Nepi et al., 2012) and preferred by bees in dual choice feeding experiments (Carter et al., 2006; Bertazzini et al., 2010). On the whole, it therefore seems that amino acid profile in rapeseed nectars differs from that hypothesized to be attractive for foraging insects. In fact, it is believed that rapeseed foraging honeybees would obtain essential amino acids and protein mainly from the large amount of pollen available from rapeseed, which shows a nutritionally balanced amino acid composition (Somerville, 2001). Moreover, recent data seem to point to a much more complex picture for the relationship between amino acid content and bee preference. For instance, the honeybee's nutritional state was found to influence the likelihood it would feed on and learn sucrose solutions containing single amino acids (Simcock et al., 2014). Moreover, the nutritional balance of essential amino acids and carbohydrates of the adult worker honeybee was shown to depend on age, and foragers were found to require a diet high in carbohydrates (essential amino acids:carbohydrates 1:250), and showed low survival rates on diets high in amino acids (Paoli et al., 2014).

Whatever the effects on insect foraging preference, the present results showed a significant intraspecific variability with respect to both sugar and amino acid content in rapeseed nectars. This is consistent with an increasing amount of evidence (Lanza et al., 1995; Wolf et al., 1999; Leiss et al., 2004; Herrera et al., 2006) that superseded an early view of a substantial constancy of nectar composition (Baker and Baker, 1977). Several environmental factors have been shown to influence nectar production. For instance, nitrogen fertilization was found to greatly influence amino acid content (Gardener and Gillman, 2001b; Gijbels et al., 2015). However, due to the adoption of uniform conditions for growth including sufficient nitrogen supply, the variability found in this study could be genetic. The characterization of phloem sap content in a subset of rapeseed genotypes shed some light on this aspect. Despite the reduced number of cultivar considered, the comparison between phloem sap and nectar content clearly showed that the latter depends on the metabolic activity of nectaries. Concerning sugars, sucrose in phloem sap is hydrolyzed into its components. If expressed as hexose equivalents, concentrations in nectar and phloem sap are quite similar. However, the glucose:fructose ratio deviates significantly from the expected 1:1, in rapeseed as well as in many other species. This discrepancy strengthens even more the role of nectaries in determining the final composition of nectar. After the hydrolysis of sucrose, the hexoses are partially cycled through various biochemical pathways before being secreted into the lumen of the nectary, and this complex metabolism could explain the different ratios observed (Brandenburg et al., 2009). With respect to free amino acids, concentrations in phloem sap and nectar show, on the contrary, a striking divergence in rapeseed, the levels in the latter being 5% of the former. Absolute concentrations of individual amino acids are much more variable than their relative ratios. In phloem sap, glutamine accounts for two thirds of total amino acid content (**Figure 8**), most likely serving as the main form for organic nitrogen transport because of a high N:C ratio. In nectar, its percent contribution halves (**Figure 6**), concomitantly with a general reduction of amino acid content. It seems likely that most amino acids are retained by nectaries to sustain their active metabolism. Most interestingly, both sugar and amino acid composition of phloem sap showed a significantly higher uniformity than that of nectar, pointing to a low variability of the former among genotypes (**Table 3**, **Figure 8**). Therefore, the variability found in nectar seems to rely mainly on a different metabolization of a relatively constant phloem sap by nectary cells.

Although floral nectar traits are important for plant reproduction, little is known about their genetic basis. Only a few studies have quantified heritable variation for nectar traits (Mitchell, 2004). Our study suggests that nectary-specific expression levels of selected enzymes may play a main role in determining the variability found among rapeseed genotypes with respect to sugar and amino acid composition of nectar. Concerning nectar volume, some experimental evidence showed an essential role of jasmonic acid levels, integrating the floral nectar secretion into the complex network of oxylipine-mediated developmental processes of plants (Radhika et al., 2010). Future work will be required to verify whether the resulting variability corresponds to a different degree of preference by foraging insects. Irrespective of this aspect, the existence of an intraspecific variability implies the possibility of breeding for the attainment of increased concentration of selected nectar components. Moreover, the identification of nectary-specific promoters in an increasing number of species and some advances in the functional genomics of nectar production (Bender et al., 2012) are opening the way toward the tailoring of nectar composition through genetic transformation.

AUTHOR CONTRIBUTIONS

MB performed most of the experiments; GF conceived and planned the research, performed a part of the experimental work, analyzed the results and wrote the paper.

ACKNOWLEDGMENTS

This work was partially supported by the University of Ferrara (*Fondo di Ateneo per la Ricerca 2012*). MB gratefully acknowledges an applied research fellowship from Spinner Global Grant, Emilia Romagna Region. The authors are indebted to Drs. Samuele Giberti and Davide Petrollino for help in nectar harvesting. The authors also thank seed companies, as listed in **Table 1**, for kindly providing *B. napus* seeds.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00288

REFERENCES

Abrol, D. P. (2007). Honeybees and rapeseed: a pollinator–plant interaction. *Adv. Bot. Res.* 45, 337–367. doi: 10.1016/S0065-2296(07)45012-1

Abrol, D. P., and Shankar, U. (2012). "Pollination in oil crops: recent advances and future strategies," in *Technological Innovations in Major World Oil Crops*, Vol. 2, ed S. K. Gupta (New York, NY: Springer), 221–267.

Antoň S., and Denisow, B. (2014). Nectar production and carbohydrate composition across floral sexual phases: contrasting patterns in two protandrous *Aconitum* species (Delphinieae, Ranunculaceae). *Flora* 209, 464–470. doi: 10.1016/j.flora.2014.07.001

Baker, H. G., and Baker, I. (1973). Amino acids in nectar and their evolutionary significance. *Nature* 241, 543–545. doi: 10.1038/241543b0

Baker, H. G., and Baker, I. (1977). Intraspecific constancy of floral nectar amino acid complements. *Bot. Gaz.* 138, 183–191. doi: 10.1086/336914

Ball, D. W. (2007). The chemical composition of honey. *J. Chem. Educ.* 84, 1643–1646. doi: 10.1021/ed084p1643

Bender, R., Klinkenberg, P., Jiang, Z., Bauer, B., Karypis, G., Nguyen, P., et al. (2012). Functional genomics of nectar production in the *Brassicaceae*. *Flora* 207, 491–496. doi: 10.1016/j.flora.2012.06.005

Bertazzini, M., Medrzycki, P., Bortolotti, L., Maistrello, L., and Forlani, G. (2010). Amino acid content and nectar choice by forager honeybees (*Apis mellifera* L.). *Amino Acids* 39, 315–318. doi: 10.1007/s00726-010-0474-x

Biancucci, M., Mattioli, R., Forlani, G., Funck, D., Costantino, P., and Trovato, M. (2015). Role of proline and GABA in sexual reproduction of angiosperms. *Front. Plant Sci.* 6:680. doi: 10.3389/fpls.2015.00680

Bosi, G. (1973). Méthode rapide pour la détermination par chromatographie en phase gazeuse des glucides du nectar: technique de prélèvement du nectar et de préparation des éthers triméthylsilyles en présence d'eau. *Apidologie* 4, 57–64. doi: 10.1051/apido:19730105

Brandenburg, A., Dell'Olivo, A., Bshary, R., and Kuhlemeier, C. (2009). The sweetest thing: advances in nectar research. *Curr. Opin. Plant Biol.* 12, 486–490. doi: 10.1016/j.pbi.2009.04.002

Carter, C., Shafir, S., Yehonatan, L., Palmer, R. G., and Thornburg, R. (2006). A novel role for proline in plant floral nectars. *Naturwissenschaften* 93, 72–79. doi: 10.1007/s00114-005-0062-1

Chambó, E. D., De Oliveira, N. T., Garcia, R. C., Duarte-Júnior, J. B., Ruvolo-Takasusuki, M. C., and Toledo, V. A. (2014). Pollination of rapeseed (*Brassica napus*) by Africanized honeybees (Hymenoptera: Apidae) on two sowing dates. *Ann. Acad. Bras. Cienc.* 86, 2087–2100. doi: 10.1590/0001-3765201420140134

Crailsheim, K., and Leonhard, B. (1997). Amino acids in honeybee worker haemolymph. *Amino Acids* 13, 141–153. doi: 10.1007/BF01373212

De Groot, A. P. (1953). Protein and amino acid requirements of the honey bee (*Apis mellifera* L.). *Phys. Comp. Oecol.* 3, 197–285.

De La Barrera, E., and Nobel, P. S. (2004). Nectar: Properties, floral aspects, and speculations on origin. *Trends Plant Sci.* 9, 65–69. doi: 10.1016/j.tplants.2003.12.003

Denisow, B., Masierowska, M., and Antòn, S. (2015). Floral nectar production and carbohydrate composition and the structure of receptacular nectaries in the invasive plant *Bunias orientalis* L. *(Brassicaceae). Protoplasma*. doi: 10.1007/s00709-015-0902-6. [Epub ahead of print].

Forlani, G., Bertazzini, M., and Giberti, S. (2014). Differential accumulation of γ-aminobutyric acid in elicited cells of two rice cultivars showing contrasting sensitivity to the blast pathogen. *Plant Biol.* 16, 1127–1132. doi: 10.1111/plb.12165

Gardener, M. C., and Gillman, M. P. (2001a). Analyzing variability in nectar amino acids: composition is less important than concentration. *J. Chem. Ecol.* 27, 2545–2558. doi: 10.1023/A:1013687701120

Gardener, M. C., and Gillman, M. P. (2001b). The effects of soil fertilizer on amino acids in the floral nectar of corncockle, *Agrostemma githago* (Caryophyllaceae). *Oikos* 92, 101–106. doi: 10.1034/j.1600-0706.2001.920112.x

Gardener, M. C., and Gillman, M. P. (2002). The taste of nectar: a neglected area of pollination ecology. *Oikos* 98, 552–557. doi: 10.1034/j.1600-0706.2002.980322.x

Giavalisco, P., Kapitza, K., Kolasa, A., Buhtz, A., and Kehr, J. (2006). Towards the proteome of *Brassica napus* phloem sap. *Proteomics* 6, 1615–9861. doi: 10.1002/pmic.200500155

Gijbels, P., Ceulemans, T., Van den Ende, W., and Honnay, O. (2015). Experimental fertilization increases amino acid content in floral nectar, fruit set and degree of selfing in the orchid *Gymnadenia conopsea*. *Oecologia* 179, 785–795. doi: 10.1007/s00442-015-3381-8

Grüter, C., and Ratnieks, F. L. W. (2011). Flower constancy in insect pollinators: Adaptive foraging behaviour or cognitive limitation? *Commun. Integr. Biol.* 4, 633–636. doi: 10.4161/cib.16972

Hansen, K., Wacht, S., Seebauer, H., and Schnuch, M. (1998). New aspects of chemoreception in flies. *Ann. N.Y. Acad. Sci.* 855, 143–147. doi: 10.1111/j.1749-6632.1998.tb10556.x

Heil, M. (2011). Nectar: Generation, regulation and ecological functions. *Trends Plant Sci.* 16, 191–200. doi: 10.1016/j.tplants.2011.01.003

Herrera, C. M., Pérez, R., and Alonso, C. (2006). Extreme intraplant variation in nectar sugar composition in an insect-pollinated perennial herb. *Am. J. Bot.* 93, 575–581. doi: 10.3732/ajb.93.4.575

Jones, D. L., Owen, A. G., and Farrar, J. F. (2002). Simple method to enable the high resolution determination of total free amino acids in soil solutions and soil extracts. *Soil Biol. Biochem.* 34, 1893–1902. doi: 10.1016/S0038-0717(02)00203-1

Kevan, P. G., Lee, H., and Shuel, R. W. (1991). Sugar ratios in nectars of varieties of canola (*Brassica napus*). *J. Apic. Res.* 30, 99–102.

Lancashire, P. D., Bleiholder, H., Langelüddecke, P., Stauss, R., Van Den Boom, T., Weber, E., et al. (1991). An uniform decimal code for growth stages of crops and weeds. *Ann. Appl. Biol.* 119, 561–601. doi: 10.1111/j.1744-7348.1991.tb04895.x

Lanza, J., Smith, G. C., Sack, S., and Cash, A. (1995). Variation in nectar volume and composition of *Impatiens capensis* at the individual, plant, and population levels. *Oecologia* 102, 113–119. doi: 10.1007/BF00333318

Leiss, K. A., Vrieling, K., and Klinkhamer, P. G. L. (2004). Heritability of nectar production in *Echium vulgare*. *Heredity* 92, 446–451. doi: 10.1038/sj.hdy.6800439

Leseigneur, C. D., Verburgt, L., and Nicolson, S. W. (2007). Whitebellied sunbirds (*Nectarinia talatala*, Nectariniidae) do not prefer artificial nectar containing amino acids. *J. Comp. Physiol. B* 177, 679–685. doi: 10.1007/s00360-007-0165-4

Lotz, C. N., and Schondube, J. E. (2006). Sugar preferences in nectar- and fruit-eating birds: Behavioral patterns and physiological causes. *Biotropica* 38, 3–15. doi: 10.1111/j.1744-7429.2006.00104.x

Lu, N.-N., Li, X.-H., Li, L., and Zhao, Z.-G. (2015). Variation of nectar production in relation to plant characteristics in protandrous *Aconitum gymnandrum*. *J. Plant Ecol.* 8, 122–129. doi: 10.1093/jpe/rtv020

Mesquida, J., Marilleau, R., Pham-Delegue, M.-H., and Renard, M. (1988). A study of rapeseed (*Brassica napus* L. var. oleifera Metzger) flower nectar secretions. *Apidologie* 19, 307–318. doi: 10.1051/apido:19880309

Mevi-Schutz, J., and Erhardt, A. (2005). Amino acids in nectar enhance butterfly fecundity: a long-awaited link. *Am. Nat.* 165, 411–420. doi: 10.1086/429150

Micheu, S., Crailsheim, K., and Leonhard, B. (2000). Importance of proline and other amino acids during honeybee flight (*Apis mellifera carnica* POLLMANN). *Amino Acids* 18, 157–175. doi: 10.1007/s007260050014

Mitchell, R. J. (2004). Heritability of nectar traits: why do we know so little? *Ecology* 85, 1527–1533. doi: 10.1890/03-0388

Mohr, N. A., and Jay, S. C. (1990). Nectar production of selected cultivars of *Brassica campestris* L. and *Brassica napus* L. *J. Apic. Res.* 29, 95–100.

Nedić, N., Maèukanović-Jocić, M., Ranèić, D., Rørslett, B., Šoštarić, I., Stevanović, Z. D., and Mladenović, M. (2013). Melliferous potential of *Brassica napus* L. subsp. *napus* (Cruciferae). *Arthropod Plant Int.* 7, 323–333. doi: 10.1007/s11829-013-9247-2

Nepi, M., Soligo, C., Nocentini, D., Abate, M., Guarnieri, M., Cai, G., et al. (2012). Amino acids and protein profile in floral nectar: much more than a simple reward. *Flora* 207, 475–481. doi: 10.1016/j.flora.2012.06.002

Nepi, M., and Stpiczyńska, M. (2008). The complexity of nectar: secretion and resorption dynamically regulate nectar features. *Naturwissenschaften* 95, 177–184. doi: 10.1007/s00114-007-0307-2

Nicolson, S. W., and Thornburg, R. W. (2007). "Nectar Chemistry," in *Nectaries and Nectar*, ed S. W. Nicolson, M. Nepi, and E. Pacini (Dordrecht: Springer), 215–264.

Paoli, P. P., Donley, D., Stabler, D., Saseendranath, A., Nicolson, S. W., Simpson, S. J., et al. (2014). Nutritional balance of essential amino acids and carbohydrates of the adult worker honeybee depends on age. *Amino Acids* 46, 1449–1458. doi: 10.1007/s00726-014-1706-2

Pierre, J., Mesquida, J., Marilleau, R., Pham-Delègue, M. H., and Renard, M. (1999). Nectar secretion in winter oilseed rape, *Brassica napus*—quantitative and qualitative variability among 71 genotypes. *Plant Breed.* 118, 471–476. doi: 10.1046/j.1439-0523.1999.00421.x

Radhika, V., Kost, C., Boland, W., and Heil, M. (2010). The role of jasmonates in floral nectar secretion. *PLoS ONE* 5:e9265. doi: 10.1371/journal.pone.0009265

Sabo, M., Potočnjak, M., Banjari, I., and Petrović, D. (2011). Pollen analysis of honeys from Varaždin County, Croatia. *Turk. J. Bot.* 35, 581–587. doi: 10.3906/bot-1009-86

Shi, J., Li, R., Zou, J., Long, Y., and Meng, J. (2011). A dynamic and complex network regulates the heterosis of yield-correlated traits in rapeseed (*Brassica napus* L.). *PLoS ONE* 6:e21645. doi: 10.1371/journal.pone.0021645

Simcock, N. K., Gray, H. E., and Wright, G. A. (2014). Single amino acids in sucrose rewards modulate feeding and associative learning in the honeybee. *J. Insect Physiol.* 69, 41–48. doi: 10.1016/j.jinsphys.2014.05.004

Somerville, D. C. (2001). *Nutritional Value of Bee Collected Pollens. RIRDC Publication No. 01/047*. Canberra: New South Wales Agriculture.

Truchado, P., Ferreres, F., Bortolotti, L., Sabatini, A. G., and Tomás-Barberán, F. A. (2008). Nectar flavonol rhamnosides are floral markers of acacia (*Robinia pseudacacia*) honey. *J. Agric. Food Chem.* 56, 8815–8824. doi: 10.1021/jf801625t

Wacht, S., Lunau, K., and Hansen, K. (2000). Chemosensory control of pollen ingestion in the hoverfly *Eristalis tenax* by labellar taste hairs. *J. Comp. Physiol. A* 186, 193–203. doi: 10.1007/s003590050019

Waddington, K. D. (1982). Honey bee foraging profitability and round dance correlates. *J. Comp. Physiol. A* 148, 297–301. doi: 10.1007/BF00679014

Westcott, L., and Nelson, D. (2001). Canola pollination: an update. *Bee World* 82, 115–129. doi: 10.1080/0005772X.2001.11099514

Williams, I., and Frank, L. (1975). Improved chemical synthesis and enzymatic assay of δ^1-pyrroline-5-carboxylic acid. *Anal. Biochem.* 64, 85–97. doi: 10.1016/0003-2697(75)90408-X

Wolf, S., Lensky, Y., and Paldi, N. (1999). Genetic variability in flower attractiveness to honeybees (*Apis mellifera* L.) within the genus *Citrullus*. *Hort. Sci.* 34, 860–863.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Bertazzini and Forlani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 March 2016
doi: 10.3389/fpls.2016.00334

Genome-Wide Single-Nucleotide Polymorphisms Discovery and High-Density Genetic Map Construction in Cauliflower Using Specific-Locus Amplified Fragment Sequencing

Zhenqing Zhao[1], Honghui Gu[1], Xiaoguang Sheng[1], Huifang Yu[1], Jiansheng Wang[1], Long Huang[2] and Dan Wang[2]*

[1] Institute of Vegetables, Zhejiang Academy of Agricultural Sciences, Hangzhou, China, [2] Biomarker Technologies Corporation, Beijing, China

OPEN ACCESS

Edited by:
Diego Rubiales, Consejo Superior de Investigaciones Científicas, Spain

Reviewed by:
Daniela Marone, Centre of Cereal Research - CREA-CER - Foggia, Italy
Michael Benjamin Kantar, University of Minneosta, USA

***Correspondence:**
Honghui Gu
guhh@mail.zaas.ac.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *18 October 2015*
Accepted: *04 March 2016*
Published: *21 March 2016*

Citation:
Zhao Z, Gu H, Sheng X, Yu H, Wang J, Huang L and Wang D (2016) Genome-Wide Single-Nucleotide Polymorphisms Discovery and High-Density Genetic Map Construction in Cauliflower Using Specific-Locus Amplified Fragment Sequencing. Front. Plant Sci. 7:334. doi: 10.3389/fpls.2016.00334

Molecular markers and genetic maps play an important role in plant genomics and breeding studies. Cauliflower is an important and distinctive vegetable; however, very few molecular resources have been reported for this species. In this study, a novel, specific-locus amplified fragment (SLAF) sequencing strategy was employed for large-scale single nucleotide polymorphism (SNP) discovery and high-density genetic map construction in a double-haploid, segregating population of cauliflower. A total of 12.47 Gb raw data containing 77.92 M pair-end reads were obtained after processing and 6815 polymorphic SLAFs between the two parents were detected. The average sequencing depths reached 52.66-fold for the female parent and 49.35-fold for the male parent. Subsequently, these polymorphic SLAFs were used to genotype the population and further filtered based on several criteria to construct a genetic linkage map of cauliflower. Finally, 1776 high-quality SLAF markers, including 2741 SNPs, constituted the linkage map with average data integrity of 95.68%. The final map spanned a total genetic length of 890.01 cM with an average marker interval of 0.50 cM, and covered 364.9 Mb of the reference genome. The markers and genetic map developed in this study could provide an important foundation not only for comparative genomics studies within *Brassica oleracea* species but also for quantitative trait loci identification and molecular breeding of cauliflower.

Keywords: cauliflower, SLAF, SNP, sequencing, genetic map

INTRODUCTION

Cauliflower (*Brassica oleracea* var. *botrytis*, 2n = 2x = 18) is an important vegetable crop worldwide. It is considered a vital source of vitamins, dietary fiber, antioxidants, and anti-carcinogenic compounds (Volden et al., 2009; Picchi et al., 2012). In 2013, global cauliflower cultivation had spread across ~1.2 million hectares, with total production of ~20.9 million metric tons (http://faostat.fao.org/). Due to the economic importance and nutritional value of cauliflower, great efforts have been taken to improve its yield and quality. Nevertheless, breeding methods used

are essentially conventional and less-effective, resulting in relative slow progress of the cauliflower breeding program (Gu et al., 2014). Modern strategies such as marker-assisted selection (MAS) are therefore necessary to accelerate cauliflower genetic improvement.

Molecular markers and genetic maps are considered an important foundation for quantitative trait loci (QTL) mapping and MAS (Kato et al., 2014; Li et al., 2014; Wu et al., 2014; Zhang et al., 2014). As one of the subspecies of *B. oleracea*, cauliflower has been employed to develop a series of segregating populations to construct several C-genome genetic maps, by crossing with other subspecies of *B. oleracea*, including kale (*B. oleracea* var. *acephala*) (Kianian and Quiros, 1992), broccoli (*B. oleracea* var. *italica*) (Li et al., 2003; Gao et al., 2007), collard (*B. oleracea* var. *acephala*) (Hu et al., 1998), and brussel sprouts (*B. oleracea* var. *gemmifera*) (Sebastian et al., 2000, 2002), and to identify several QTLs involved in common traits such as flowering time (Kianian and Quiros, 1992), glucosinolate profile (Gao et al., 2007), and leaf traits (Sebastian et al., 2002). However, because of the organ specificity of cauliflower curd, inter-subspecies populations are unsuitable for QTL analysis of many curd-specific traits, which are most important for cauliflower breeding. Hence, a cauliflower × cauliflower based population has extensive potential for marker discovery, genetic mapping, and QTL analysis.

Several molecular markers such as restriction fragment length polymorphism (RFLP), randomly-amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), sequence-related amplified polymorphism (SRAP), and simple sequence repeats (SSR) have been widely used in plant genetic research. However, the utilization of these markers has always been limited by the higher time and cost requirements as well as limited marker resource (Zhao et al., 2014a). The advent of massive, parallel, next-generation sequencing (NGS) technologies have accelerated and simplified the identification of sequence variants, enabling large-scale single-nucleotide polymorphism (SNP) discovery throughout the genome (Zhou et al., 2014). Considering that whole-genome deep re-sequencing is still expensive and usually not necessary (Davey et al., 2011; Wei et al., 2014), several simplified and cost-effective methods for SNP discovery and high-throughput genotyping have been developed, such as reduced representation library (RRL) sequencing (Van Tassell et al., 2008), restriction-site associated DNA sequencing (RAD-seq; Miller et al., 2007), two-enzyme genotyping by sequencing (GBS) (Poland et al., 2012), and sequence-based genotyping (SBG) (Truong et al., 2012). More recently, specific-locus amplified fragments sequencing (SLAF-seq) was developed as a streamlined RRL sequencing approach for high-resolution de novo SNP discovery and genotyping (Sun X. et al., 2013). A high density genetic map including 1233 high-quality markers has been developed using this strategy on a sesame F_2 population (Zhang et al., 2013). This study showed that SLAF sequencing was a powerful high-throughput technique for plant genome research. To date, this strategy has also been successfully applied in several other species including soybean (Li et al., 2014; Qi et al., 2014), cucumber (Wei et al., 2014; Xu et al., 2014), tea plant (Ma et al., 2015), and grape (Guo et al., 2015).

In this study, we generated a double-haploid (DH) population derived from a cross between two different types of cauliflower common in production, including an advanced inbred line of traditional compact-curd cauliflower and a DH line of loose-curd cauliflower. Based on this cauliflower × cauliflower population, SLAF-seq was then employed to detect large-scale SNPs and construct a high-density genetic map that could be used to provide a platform for future QTL mapping and MAS.

MATERIALS AND METHODS

Mapping Population Development and Genomic DNA Isolation

An advanced inbred line of compact-curd cauliflower "4305" (F_8) and a DH line of loose-curd cauliflower "ZN198" were used to develop the DH mapping population (Figure S1). There are significant morphological differences between the two homozygous lines, especially for several agronomically important traits including curd size, curd weight, curd shape, and curding-time. A microspore culture protocol as previously described by Gu et al. (2014) was used to produce regenerated plants from a single F_1 plant of the cross "4305" × "ZN198." The ploidy level of all the regenerated plants achieved was estimated using an FCM Ploidy Analyzer (Partec GmbH, Germany) and only diploids were selected to construct the mapping population. Parents and DH lines were planted in the experiment field of Zhejiang Academy of Agricultural Sciences in Hangzhou, China and were preserved for long-term utilization by artificial selfing.

Young leaves were collected and genomic DNA was isolated according to a modified version of the cetyltrimethyl ammonium bromide (CTAB) procedure (Doyle and Doyle, 1986). DNA concentration and quality were detected using an ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA) and electrophoresis on 1.0% agarose gel with a standard lambda DNA.

SLAF Library Preparation and High Throughput Sequencing

A SLAF-seq strategy was used, as previously described by Sun X. et al. (2013), with modifications. First, reference genome of *B. oleracea* (cabbage, *B. oleracea* var. *capitata*, http://www.ocri-genomics.org/bolbase/, Liu S. et al., 2014) was used to design marker discovery experiments by simulating *in silico*, the number of markers produced by different enzymes. A SLAF pilot experiment was performed to determine the optimized enzymes and restriction fragment size, while the SLAF library was conducted based on the pre-designed scheme. Subsequently, the genomic DNA (2 μg) was digested with 3.6 units *Rsa*I (New England Biolabs, NEB, USA) in 20 μl volume containing 1 × NEB buffer at 37°C for 2 h, and then a single nucleotide (A) was added by using 6 units Klenow Fragment ($3'\rightarrow5'$ exo$^-$) (NEB) and 10 nmol dATP at 37°C for 1 h. Duplex tag-labeled sequencing adapters (PAGE-purified, Life Technologies, USA) were then ligated to the A-tailed fragments using T4 DNA ligase (NEB) by incubating overnight at 16°C, then 65°C for 20 min to heat deactivate the T4 ligase. Polymerase chain reaction (PCR) was performed in a 100 μl final volume reaction mixture,

which contained diluted restriction-ligation DNA samples, PCR primers (forward sequence: 5′-AATGATACGGCGACCACCGA-3′, reverse sequence: 5′-CAAGCAGAAGACGGCATACG-3′), dNTP, $MgCl_2$ and Q5® High-Fidelity DNA Polymerase (PAGE-purified, Life Technologies). The PCR cycles were 98°C for 3 min, 18 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 30 s, followed by an extension step of 5 min at 72°C before storage at 4°C. The amplification products were purified using Agencourt AMPure XP beads (Beckman Coulter, High Wycombe, UK) and pooled, followed by separation on 2% agarose gel electrophoresis. Fragments ranging from 244 to 314 bp (with indices and adaptors) in size were gel-purified using a QIAquick gel extraction kit (Qiagen, Hilden, Germany) and diluted for pair-end sequencing (125 bp at each end) using an Illumina HiSeq 2500 system (Illumina, Inc.; San Diego, CA, USA) at Beijing Biomarker Technologies Corporation, according to the manufacturer recommendations.

Sequence Data Grouping and Genotyping

SLAF marker identification and genotyping were performed as previously described by Sun X. et al. (2013) and Zhang et al. (2015). First, raw reads were demultiplexed to individuals according to the barcode sequences. The reads with quality scores <Q30 (a quality score of 30; indicating 0.1% chance of an error, and thus 99.9% confidence) were filtered out. After the barcodes and the terminal 5-bp positions were trimmed, high-quality reads from the same sample were mapped onto the reference genome sequence using SOAP software (Li et al., 2008). Subsequently, sequences locating at the same position with over 95% identity were grouped into one SLAF locus. SNP of each locus were firstly detected between parents, and SLAFs with less than three SNPs were used to define alleles. As cauliflower is diploid, only SLAFs with two to four alleles were identified as polymorphic and considered potential markers. Each polymorphic SLAF marker was then classified into eight segregation patterns in population (ab × cd, ef × eg, hk × hk, lm × ll, nn × np, aa × bb, ab × cc, and cc × ab) as previously described by Zhang et al. (2013). Since the DH mapping population used here was derived from two homozygous lines, only the SLAF markers showing aa × bb segregation pattern were used for map construction. Genotype scoring was then performed using a Bayesian approach to further ensure the genotyping quality, and high-quality SLAF markers for the genetic mapping were filtered by criteria as previously described by Sun X. et al. (2013).

Genetic Map Construction

The genetic map was constructed as previously described by Zhang et al. (2015). Marker loci were allocated primarily into nine linkage groups (LGs) based on their locations on the reference genome. Markers with the modified logarithm of odds (MLOD) score <5 were filtered to further confirm the marker robustness. A newly developed High Map strategy (Liu D. et al., 2014) was applied to order the SLAF markers and correct genotyping errors within LGs. MSTmap algorithm was used to order SLAFs markers (Wu et al., 2008) and the SMOOTH algorithm (Hans et al., 2005) was used to correct genotyping errors following marker ordering. Map distances were estimated using the Kosambi mapping function (Kosambi, 1943). Since all the mapped markers have been mapped onto the reference genome of *B. oleracea* by SOAP software (Li et al., 2008), the collinearity of physical map and genetic map were visualized by RSCRIPT language following the tutorial introduction.

RESULTS

The Mapping Population

In total, 136 regenerated plants with different genotypes were obtained from "4305" × "ZN198" F_1 by microspore culture. Several ploidy levels were detected in these plants, including haploids, diploids, polyploids, and chimeras (Figure S2). Finally, 79 diploid plants showing a spontaneously doubling ratio of 58.1% were obtained and used as the mapping population.

SLAF Sequencing and Genotyping

A total of 12.47 Gb raw data containing 77.92 M pair-end reads were generated from the high-throughput sequencing (Data has been submitted to National Center of Biotechnology Information, the BioProject ID was PRJNA307521), with a GC (guanine-cytosine) content of 38.12%. The rate of high-quality reads with quality scores >30 reached 91.07%. After read clustering and filtering, 81,311 high-quality SLAFs with even distribution throughout the genome were identified (**Figure 1A**; **Table 1**; Table S1). The average sequencing depths of these SLAFs were 52.66-fold for female inbred line (4305), 49.35-fold for male DH line (ZN198), and 4.71-fold for each progeny of the DH population (**Figure 2**). Based on the allele number and sequence difference, the 81,311 SLAFs were grouped into three types including polymorphic, non-polymorphic, and repetitive. 6815 polymorphic SLAFs out of all SLAFs were obtained, showing a polymorphism rate of 8.38% (**Figure 1B**; Table S2). Interestingly, the polymorphism rate of SLAFs developed on each chromosome was quite different. In the current study, SLAFs on chromosome C01, C07, and C08 showed significantly higher polymorphism rate than the others (**Table 2**).

Subsequently, all these polymorphic SLAFs were genotyped separately for parents and population individuals. A total of 6568 polymorphic SLAFs from total 6815 were successfully encoded, of which 4736 SLAFs were classified as the expected aa × bb segregation pattern, following the genotype encoding rule (**Figure 3**). These 4736 SLAFs were further filtered based on the criteria considering segregation distortion, sequencing depth, and data integrity. Finally, 1776 high-quality markers, with the average sequencing depths of 63.05-fold in the female parent, 59.75-fold in the male parent, and 4.50-fold in each DH individual, were employed to construct the genetic map (**Figure 1C**; **Table 1**).

Main Characteristics of the Genetic Map

All 1776 high-quality markers were successfully assigned onto 9 LGs according to their locations on the reference genome and the MLOD scores with other markers (at least one MLOD score >5). The average data integrity of these 1776 SLAF markers reached 95.68%. The final genetic map was constructed,

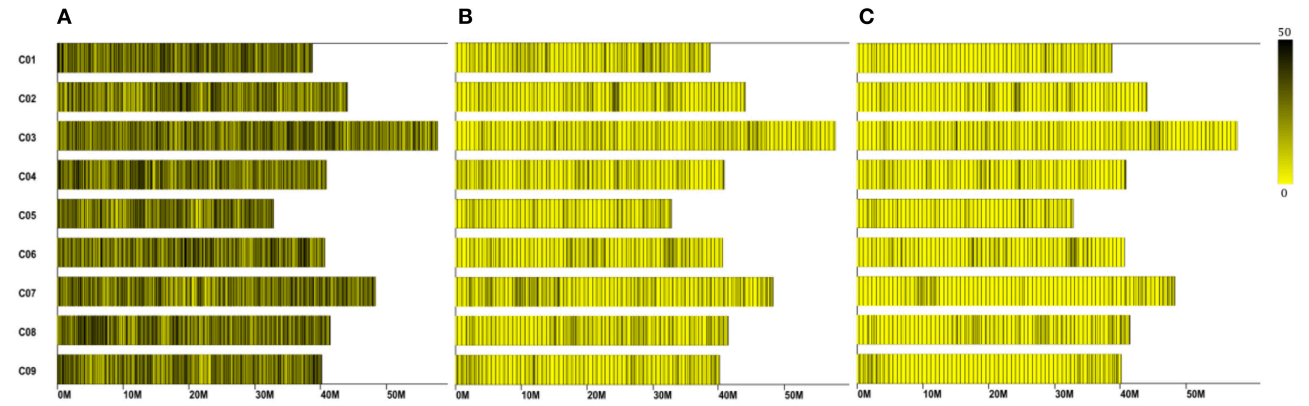


FIGURE 1 | Distribution of the total SLAFs (A) polymorphic SLAFs (B) and mapped SLAF markers (C) on the reference genome. The x-axis represents chromosome length and the y-axis indicates chromosome code. Each yellow bar stands for a chromosome, and deeper color from yellow to black means more SLAFs on the corresponding location.

TABLE 1 | SLAF-seq data summary for the mapping population.

TOTAL READS	
No. of reads	77.92 M
HIGH QUALITY SLAFs	
No. SLAFs	81,311
Average SLAF depth	6.33
Average depth in parents	51.01
Average depth in population individuals	4.71
POLYMORPHIC SLAFs	
No. of polymorphic SLAFs	6815
Average depth in parents	57.00
Average depth in population individuals	4.52
HIGH-QUALITY SLAF MARKERS	
No. of high-quality SLAF markers	1776
No. of SNPs	2741
Average depth in parents	61.40
Average depth in population individuals	4.50

TABLE 2 | Polymorphism rate of SLAFs developed on each chromosome.

Chromosome ID	Total SLAF	Polymorphic SLAF	Polymorphism rate %
C01	8856	857	9.68
C02	8589	676	7.87
C03	12,049	898	7.45
C04	8779	654	7.45
C05	6918	453	6.55
C06	8797	726	8.25
C07	9892	1053	10.64
C08	9096	845	9.29
C09	8335	653	7.83
Total	81,311	6815	8.38

following linkage analysis for each of the 9 LGs, which was designated according to the corresponding chromosome number of the reference genome (**Figure 4**; **Table 3**). A total of 1776 mapped SLAF markers containing 2741 SNPs spanned a total genetic length of 890.01 cM, with an average marker interval of 0.50 cM and covered 364.9 Mb of the reference genome. The genetic length, marker number, and average marker interval of single LG ranged from 41.90 (C01) to 200.92 (C06), 144 (C01) to 277 (C03), and 0.23 (C02) to 0.87 cM (C06), respectively. The max gap was 12.94 cM, located on C07. Detailed data of the genetic map and markers are presented in Table S3.

Visualization and Evaluation of the Genetic Map

The mapped markers were anchored on the reference genome and the correlation of genetic and physical position is shown in Figure S4. Generally, a sufficient genome coverage and accurate genetic location of markers were revealed by the consecutive curves. The genetic arrangements of most markers were also considered to coincide with their physical direction based on the falling trend of the curve. However, a significant inversion can be observed intuitively on pseudo-chromosome 4. The detailed collinearity and marker location are described in Figure S3; Table S3.

Haplotype maps were generated for each DH individual with two parents as control using all the mapped markers in order to detect double recombination and deletion, thus to reflect potential genotyping and marker-order errors (West et al., 2006). In this study, there was no double recombination and deletion found in any linkage group (Figure S5). In addition, the map quality was also evaluated by heat maps, which could intuitively display recombination relationships among markers within each single LG. Pair-wise recombination rates could be visualized by different color levels from yellow to purple. Here, yellow color generally showed diagonal distribution in the heat map for each of the nine LGs (Figure S6), indicating that the mapped marker had been correctly ordered.

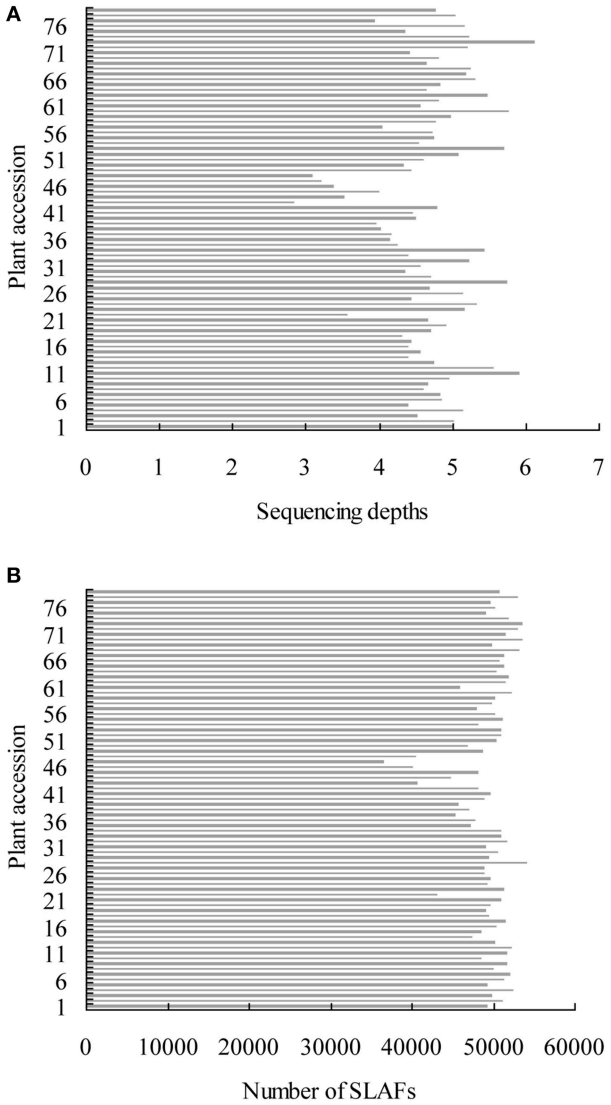


FIGURE 2 | Sequencing depths and number of SLAFs of all the individuals of mapping population. The x-axis indicates the average depths **(A)** and number of SLAFs **(B)** the y-axis represents individual accessions.

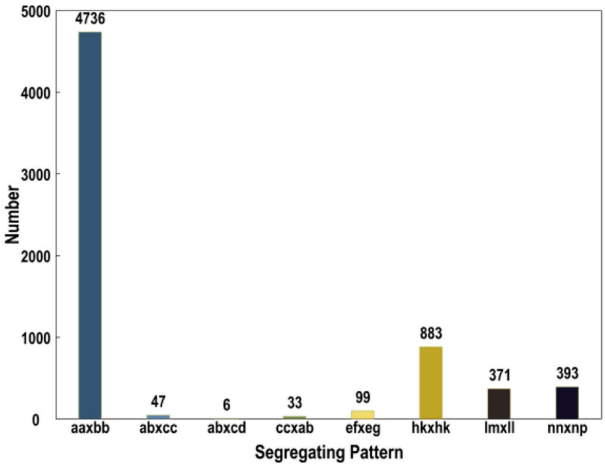


FIGURE 3 | Number of SLAF markers in each segregation patterns.

DISCUSSION

Due to the genome-wide abundance and high-throughput nature, SNP markers are playing an increasingly important role in plant research, such as genetic map construction, novel gene discovery, evolutionary analysis, and MAS within breeding programs (Liu et al., 2012). SLAF-seq is a newly reported enhanced RRL sequencing solution for large-scale SNP discovery and genotyping (Sun X. et al., 2013). It can produce large amounts of sequence-based information and handle any density distribution throughout the whole genome (Chen et al., 2013). Moreover, compared with another widely used, NGS-based method, RAD-seq, SLAF-seq shows higher reproducibility (Qi et al., 2014; Xu et al., 2014), and locus discrimination efficiency (Wei et al., 2014) due to its paired-ends sequencing strategy and longer read length (30–50 bp), respectively. In this study, we used this cost-effective strategy to develop and sequence a total of 81,311 SLAFs. Among these, 6815 SLAFs showing polymorphism between two parents ("4305" and "ZN198") were identified with a polymorphic rate of 8.38%. Although the rate was much lower than that previously obtained by whole-genome re-sequencing in cabbage (Wang et al., 2012), the markers developed herein covered all pseudo-chromosomes and were evenly distributed throughout the reference genome (**Figure 1B**). In addition, the reads quality score of > 30 and an average sequencing depth of 57-fold for parents (**Table 1**) ensured high genotyping accuracy. Therefore, this set of SLAFs has great potential for use in both genomic study and breeding application of cauliflower. Additionally, our study further demonstrated that SLAF-seq is an efficient strategy for genome-wide SNP identification by sampling and sequencing a reduced set of representative genome regions instead of the whole genome.

Long-term domestication and selective breeding have resulted in abundant variations within *B. oleracea* species, hence several morphologically different subspecies currently available worldwide, such as cabbage, brussel sprouts, cauliflower, and broccoli (Kennard et al., 1994). In order to make the most of genetic variation within *B. oleracea*, traditional genomic researches usually employ different subspecies as parental materials to develop markers, construct genetic maps, and identify QTLs (Wang et al., 2012). However, these advances have been not comprehensive enough to support the genetic improvement of some specific organ, like head of cabbage or curd of cauliflower. For this reason, increasing attention has been given to genomic study based on intra-subspecies populations. For instance, Wang et al. (2012) developed more than 5000 SSR and SNP markers using two cabbage inbred lines and constructed a high-density genetic map including 1227 markers. Based on another cabbage × cabbage population, Lv et al. (2014) identified 707 InDels (insertion–deletions) and detected 13 reliable QTLs associated with five important heading traits. Among the *B. oleracea* varieties, cauliflower is the only

TABLE 3 | Basic characteristics of the nine LGs.

LG ID	Total SLAFs	Total SNPs	Size (cM)	Physical length (Mb)	Average Marker Interval (cM)	Gap ≤ 5 (%)
C01	144	219	41.90	36.6	0.29	98.60
C02	199	291	45.61	41.7	0.23	99.49
C03	277	420	89.97	53.9	0.32	99.64
C04	213	320	158.70	39.7	0.75	96.23
C05	128	196	90.01	31.3	0.70	97.64
C06	232	355	200.92	37.1	0.87	93.07
C07	164	267	105.77	44.5	0.64	95.09
C08	274	439	82.31	40.6	0.30	100.00
C09	145	234	74.82	39.3	0.52	97.22
Total	1776	2741	890.01	364.9	0.50	98.60

"Gap ≤ 5" indicates the percentage of gaps in which the distance between adjacent markers was smaller than 5 cM.

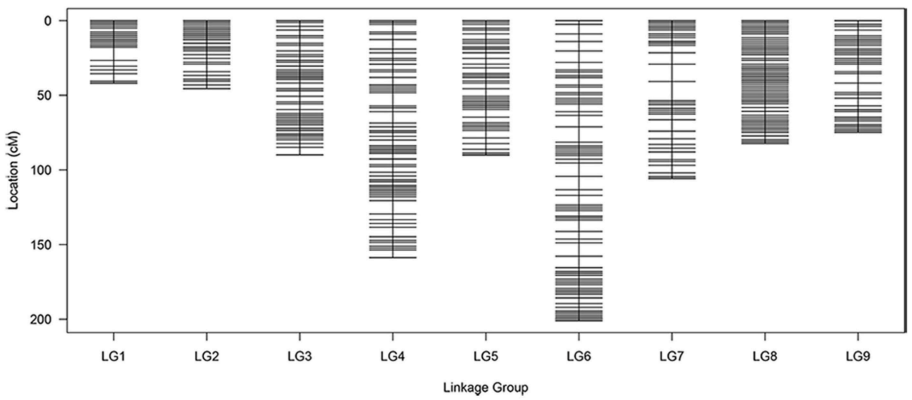


FIGURE 4 | Distribution of SLAF markers on eight linkage groups of cauliflower. A black bar indicates a SLAF marker. The x-axis represents linkage group number and the y-axis indicates genetic distance.

plant where the immature inflorescence shows hypertrophic structure (curd). Therefore, the improvement of curd-specific traits like weight, size, shape, color, and content of nutritional compounds are important goals for the cauliflower breeding program. In this study, we used two different types of cauliflower, commonly produced in China, including an advanced inbred line of traditional compact-curd cauliflower and a DH line of loose-curd cauliflower, to identify SNPs and construct a genetic map. Firstly, the genetic base difference between loose-curd cauliflower and compact-curd cauliflower (Zhao et al., 2014b) would ensure acceptable marker polymorphism. Secondly, loose-curd cauliflower is now rapidly widespread and becoming a main cultivated type of cauliflower in China (Zhao et al., 2012). It shows varied agronomic characteristics (Zhao et al., 2013) and nutritional compounds (Gu et al., 2015) with a compact curd. The cauliflower-based SNP information should be helpful to enhance the future breeding program. In our lab, QTL mapping that aims to uncover the genetic factors controlling a series of curd-specific traits and to identify related markers based on the current study is now underway. In addition, loose-curd cauliflower originating from a portion of the genetic variation of domesticated cauliflower has undergone founder effects and intense selection for locally favorable curd characteristics (Zhao et al., 2014b). The significantly higher polymorphism percentage of SLAFs developed on C01, C07, and C08 between two parents suggest that these chromosomes may carry more genes/QTLs related to curd traits that have been strongly selected by breeders and cultivators. An interesting future task is to more thoroughly dissect the domestication of loose-curd cauliflower by using more abundant germplasm of both type of this crop.

Due to the limited genetic base within cauliflower (Zhao et al., 2014b), there has been very few genetic mapping studies using cauliflower based on intra-subspecies cross. The only case to our knowledge is a cauliflower × cauliflower-based map spanning a genetic length of 668.4 cM with 234 AFLP and 21 NBS (nucleotide binding site) markers (Gu et al., 2007). However, the limited marker numbers and density confined its further application. Besides, the markers in this map were difficult to be anchored onto the reference genome or transferred to other genetic maps. For all *B. oleracea* sub-species, the genetic map with highest marker density developed to date was based on *B. oleracea* Genome Sequencing Project (BrGSP) and whole-genome re-sequencing in cabbage. In this case, the map comprised 1227 markers, with an average interval of 0.98 cM,

and was applied to anchor assembled scaffolds onto pseudo-chromosomes in BrGSP (Wang et al., 2012). In contrast, the SLAF-seq-based map in the present study contained 1776 SLAF markers including 2741 SNPs and covered 364.9 Mb of the cabbage genome. The marker quantity and resolution herein were significantly improved. However, many markers on the current map were noted to be highly clustered, some of them even located on the same locus, although their corresponding physical positions were quite different (Table S3). Therefore, despite the average marker interval being as short as 0.50 cM, there were still several obvious gaps on some regions. The similar phenomenon has been reported previously as well (Wang et al., 2012; Qi et al., 2014; Cai et al., 2015), indicating that the genetic differences or recombination between mapping parents on some genome regions were inadequate (Sun L. et al., 2013). In any case, the cluster markers may be separated distinctly if they are used for a different or broader population. For single LG, we also noted the discrepancy between the genetic length and corresponding pseudo-chromosome length of *B. oleracea* draft genome. For example, C06 is the longest LG in our map with a genetic length of 200.92 cM, but it is a relative short pseudo-chromosome in reference genome draft (http://www.ocri-genomics.org/bolbase/). This case, though common, could represent the extensive regions of increased/decreased recombination or could potentially indicate areas of the draft sequence that might require further refinement.

Comparative analyses of *Brassica* accessions have demonstrated that both genes relocation and sequence polymorphisms between species are common in the *Brassica* genome (Hu et al., 1998; Sebastian et al., 2000). Here, we show that the genome of cabbage and cauliflower are highly collinear, with macro collinearity punctuated by rearrangements mainly involving translocations and inversions. The majority of the observed rearrangements involved very short distances, but an apparent inversion spanning distance of more than 20 cM was noted at the central section of linkage group 04 between Marker13078 and Marker17566 (Figure S3; Table S3). Similar violation of genetic collinearity have also been identified within both intraspecific and interspecific level in other *Brassica* species, which may be responsible for the high degree of morphological polymorphism among these species (Sebastian et al., 2000; Bancroft et al., 2011). Since the cabbage and cauliflower subspecies shared a polyphyletic origin in primitive *B. oleracea* populations (Truco et al., 1996), our results imply that the inverted region seem to have been involved in the divergence of cauliflower and cabbage. Although additional genomic data from other varieties/lines are necessary to trace the genome evolution, comparative information about the collinearity between these two closely related subspecies has important implications for further marker and gene identification in cauliflower through the use of sequence data from cabbage as a genomic model.

In conclusion, we demonstrated the application of SLAF-seq strategy in cauliflower for large-scale SNP discovery and high-density genetic map construction. The parents used in this study are not only representatives of cultivated cauliflower, but also elite lines in our breeding program. Therefore, the genome-wide SNPs identified and the linkage map developed in this study will provide an important foundation not only for QTL identification and map-based cloning, especially of curd development-related traits, but also for MAS within cauliflower breeding programs.

AUTHOR CONTRIBUTIONS

ZZ and HG conceived and performed experiments and wrote the manuscript; XS, HY, and JW were involved in data analysis; LH and DW performed the SLAF library construction and sequencing. All the authors have commented, read and approved the final manuscript.

FUNDING

This project is supported by the National Natural Science Foundation of China (grant number, 31501768); the Research project in Zhejiang Province Science and Technology department (grant number, 2012C12903-3-4, 2016C32102); the Research project in Zhejiang Academy of Agricultural Sciences (grant number, 2014CX031, 2015R23R08E05).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00334

REFERENCES

Bancroft, I., Morgan, C., Fraser, F., Higgins, J., Wells, R., Clissold, L., et al. (2011). Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nat. Biotechnol.* 29, 762–766. doi: 10.1038/nbt.1926

Cai, C., Cheng, F., Wu, J., Zhong, Y., and Liu, G. (2015). The first high-density genetic map construction in tree peony (*Paeonia* Sect. *Moutan)* using genotyping by specific-locus amplified fragment sequencing. *PLoS ONE* 10:e0128584. doi: 10.1371/journal.pone.0128584

Chen, S., Huang, Z., Dai, Y., Qin, S., Gao, Y., Zhang, L., et al. (2013). The development of 7E Chromosome-specific molecular markers for *Thinopyrum elongatum* based on SLAF-seq technology. *PLoS ONE* 8:e65122. doi: 10.1371/journal.pone.0065122

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

Doyle, J., and Doyle, J. (1986). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Gao, M., Li, G., Yang, B., Qiu, D., Farnham, M., and Quiros, C. (2007). High-density *Brassica oleracea* linkage map: identification of useful new linkages. *Theor. Appl. Genet.* 115, 277–287. doi: 10.1007/s00122-007-0568-3

Gu, H., Wang, J., Zhao, Z., Sheng, X., Yu, H., and Huang, W. (2015). Characterization of the appearance, health-promoting compounds, and antioxidant capacity of the florets of the loose-curd cauliflower. *Int. J. Food Properties* 18, 392–402. doi: 10.1080/10942912.2013.831445

Gu, H., Zhao, Z., Sheng, X., Yu, H., and Wang, J. (2014). Efficient doubled haploid production in microspore culture of loose-curd cauliflower (*Brassica oleracea* var. *botrytis*). *Euphytica* 195, 467–475. doi: 10.1007/s10681-013-1008-x

Gu, Y., Zhao, Q. C., and Sun, D. L. (2007). Construction of genetic linkage map and localization of NBS-LRR like resistance gene analogues in cauliflower (*Brassica oleracea* var. *botrytis*). *Hereditas* 29, 751–757. doi: 10.1360/yc-007-0751

Guo, Y., Shi, G., Liu, Z., Zhao, Y., Yang, X., Zhu, J., et al. (2015). Using specific length amplified fragment sequencing to construct the high-density genetic map for *Vitis* (*Vitis vinifera* L.× *Vitis amurensis* Rupr.). *Front. Plant Sci.* 6:393. doi: 10.3389/fpls.2015.00393

Hans, V. O., Stam, P., Visser, R. G. F., and Eck, H. J. V. (2005). Smooth: a statistical method for successful removal of genotyping errors from high-density genetic linkage data. *Theor. Appl. Genet.* 112, 187–194. doi: 10.1007/s00122-005-0124-y

Hu, J., Sadowski, J., Osborn, T. C., Landry, B. S., and Quiros, C. F. (1998). Linkage group alignment from four independent (*Brassica oleracea*) RFLP maps. *Genome* 41, 226–235. doi: 10.1139/g98-007

Kato, S., Sayama, T., Fujii, K., Yumoto, S., Kono, Y., Hwang, T. Y., et al. (2014). A major and stable QTL associated with seed weight in soybean across multiple environments and genetic backgrounds. *Theor. Appl. Genet.* 127, 1365–1374. doi: 10.1007/s00122-014-2304-0

Kennard, W. C., Slocum, M. K., Figdore, S. S., and Osborn, T. C. (1994). Genetic analysis of morphological variation in *Brassica Oleracea* using molecular markers. *Theor. Appl. Genet.* 87, 721–732. doi: 10.1007/BF00222898

Kianian, S. F., and Quiros, C. F. (1992). Generation of a *Brassica oleracea* composite RFLP map: linkage arrangements among various populations and evolutionary implications. *Theor. Appl. Genet.* 84, 544–554. doi: 10.1007/bf00224150

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Li, B., Ling, T., Zhang, J., Long, H., Han, F., Yan, S., et al. (2014). Construction of a high-density genetic map based on large-scale markers developed by specific length amplified fragment sequencing (SLAF-seq) and its application to QTL analysis for isoflavone content in *Glycine max*. *BMC Genomics* 15:1086. doi: 10.1186/1471-2164-15-1086

Li, G., Gao, M., Yang, B., and Quiros, C. F. (2003). Gene to gene alignment between the *Brassica* and *Arabidopsis* genomes by transcriptional mapping. *Theor. Appl. Genet.* 107, 168–180. doi: 10.1007/s00122-003-1236-x

Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: short oligonucleotide alignment program. *Bioinformatics* 24, 713–714. doi: 10.1093/bioinformatics/btn025

Liu, D., Ma, C., Hong, W., Huang, L., Liu, M., Liu, H., et al. (2014). Construction and analysis of high-density linkage map using high-throughput sequencing data. *PLoS ONE* 9:e98855. doi: 10.1371/journal.pone.0098855

Liu, J., Huang, S., Sun, M., Liu, S., Liu, Y., Wang, W., et al. (2012). An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods* 8:34. doi: 10.1186/1746-4811-8-34

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930. doi: 10.1038/ncomms4930

Lv, H., Wang, Q., Zhang, Y., Yang, L., Fang, Z., Wang, X., et al. (2014). Linkage map construction using InDel and SSR markers and QTL analysis of heading traits in *Brassica oleracea* var. *capitata L. Mol. Breed.* 34, 1–12. doi: 10.1007/s11032-014-0019-1

Ma, J. Q., Huang, L., Ma, C. L., Jin, J. Q., Li, C. F., Wang, R. K., et al. (2015). Large-scale SNP discovery and genotyping for constructing a high-density genetic map of tea plant using specific-locus amplified fragment sequencing (SLAF-seq). *PLoS ONE* 10:e0128798. doi: 10.1371/journal.pone.0128798

Miller, M. R., Dunham, J. P., Amores, A., Cresko, W. A., and Johnson, E. A. (2007). Rapid and cost effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.* 17, 240–248. doi: 10.1101/gr.5681207

Picchi, V., Migliori, C., Scalzo, R. L., Campanelli, G., Ferrari, V., and Cesare, L. F. D. (2012). Phytochemical content in organic and conventionally grown Italian cauliflower. *Food Chem.* 130, 501–509. doi: 10.1016/j.foodchem.2011.07.036

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7:e32253. doi: 10.1371/journal.pone.0032253

Qi, Z., Huang, L., Zhu, R., Xin, D., Liu, C., Han, X., et al. (2014). A high-density genetic map for soybean based on specific length amplified fragment sequencing. *PLoS ONE* 9:e104871. doi: 10.1371/journal.pone.0104871

Sebastian, R. L., Howell, E. C., King, G. J., Marshall, D. F., and Kearsey, M. J. (2000). An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled-haploid mapping populations. *Theor. Appl. Genet.* 100, 75–81. doi: 10.1007/s001220050011

Sebastian, R. L., Kearsey, M. J., and King, G. J. (2002). Identification of quantitaive trait loci controlling developmental characteristics of *Brassica oleracea* L. *Theor. Appl. Genet.* 104, 601–609. doi: 10.1007/s001220100743

Sun, L., Yang, W., Zhang, Q., Cheng, T., Pan, H., Xu, Z., et al. (2013). Genome-wide characterization and linkage mapping of simple sequence repeats in mei (*Prunus mume* Sieb. *et Zucc.). PLoS ONE* 8:e59562. doi: 10.1371/journal.pone.0059562

Sun, X., Liu, D., Zhang, X., Li, W., Liu, H., Hong, W., et al. (2013). SLAF-seq: an efficient method of large-scale de novo SNP discovery and genotyping using high-throughput sequencing. *PLoS ONE* 8:e58700. doi: 10.1371/journal.pone.0058700

Truco, M. J., Hu, J., Sadowski, J., and Quiros, C. F. (1996). Inter-and intra-genomic homology of the *Brassica* genomes: implications for their origin and evolution. *Theor. Appl. Genet.* 93, 1225–3123. doi: 10.1007/BF00223454

Truong, H. T., Ramos, A. M., Yalcin, F., De, R. M., Hj, V. D. P., Huvenaars, K. H., et al. (2012). Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. *PLoS ONE* 7:e37565. doi: 10.1371/journal.pone.0037565

Van Tassell, C., Smith, T., Matukumalli, L., Taylor, J., Schnabel, R., Lawley, C., et al. (2008). SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat. Methods* 5, 247–252. doi: 10.1038/nmeth.1185

Volden, J., Bengtsson, G. B., and Wicklund, T. (2009). Glucosinolates; L-ascorbic acid, total phenols, anthocyanins, antioxidant capacities and colour in cauliflower (*Brassica oleracea* L. ssp. *botrytis);* effects of long-term freezer storage. *Food Chem.* 112, 967–976. doi: 10.1016/j.foodchem.2008.07.018

Wang, W., Huang, S., Liu, Y., Fang, Z., Yang, L., Hua, W., et al. (2012). Construction and analysis of a high-density genetic linkage map in cabbage (*Brassica oleracea* L. var. capitata). *BMC Genomics* 13:523. doi: 10.1186/1471-2164-13-523

Wei, Q., Wang, Y., Qin, X., Zhang, Y., Zhang, Z., Jing, W., et al. (2014). An SNP-based saturated genetic map and QTL analysis of fruit-related traits in cucumber using specific-length amplified fragment (SLAF) sequencing. *BMC Genomics* 15:1158. doi: 10.1186/1471-2164-15-1158

West, M. A., Van, L. H., Kozik, A., Kliebenstein, D. J., Doerge, R. W., St Clair, D. A., et al. (2006). High-density haplotyping with microarray-based expression and single feature polymorphism markers in Arabidopsis. *Genome Res.* 16, 787–795. doi: 10.1101/gr.5011206

Wu, J., Li, L. T., Li, M., Khan, M. A., Li, X. G., Chen, H., et al. (2014). High-density genetic linkage map construction and identification of fruit-related QTLs in pear using SNP and SSR markers. *J. Exp. Bot.* 65, 5771–5781. doi: 10.1093/jxb/eru311

Wu, Y., Bhat, P. R., Close, T. J., and Lonardi, S. (2008). Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet.* 4:e1000212. doi: 10.1371/journal.pgen.1000212

Xu, X., Xu, R., Zhu, B., Yu, T., Qu, W., Lu, L., et al. (2014). A high-density genetic map of cucumber derived from Specific Length Amplified Fragment sequencing (SLAF-seq). *Front. Plant Sci.* 5:768. doi: 10.3389/fpls.2014.00768

Zhang, D., Hua, Y., Wang, X., Zhao, H., Shi, L., and Xu, F. (2014). A high-density genetic map identifies a novel major QTL for boron efficiency in oilseed rape (*Brassica napus* L.). *PLoS ONE* 9:e112089. doi: 10.1371/journal.pone.0112089

Zhang, J., Zhang, Q., Cheng, T., Yang, W., Pan, H., Zhong, J., et al. (2015). High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc). *DNA Res.* 22, 1–9. doi: 10.1093/dnares/dsv003

Zhang, Y., Wang, L., Xin, H., Li, D., Ma, C., Ding, X., et al. (2013). Construction of a high-density genetic map for sesame based on large scale marker development by specific length amplified fragment (SLAF) sequencing. *BMC Plant Biol.* 13:141. doi: 10.1186/1471-2229-13-141

Zhao, Z., Gu, H., Sheng, X., Yu, H., Wang, J., Cao, J., et al. (2012). Genetic purity testing of loose-curd cauliflower hybrids using SSR markers and grow out test. *Seed Sci. Technol.* 40, 209–214. doi: 10.15258/sst.2012.40.2.06

Zhao, Z., Gu, H., Sheng, X., Yu, H., Wang, J., and Cao, J. (2014a). Advances and applications in crop quantitative trait loci. *J. Nucl. Agric. Sci.* 28, 1615–1624. doi: 10.11869/j.issn.100-8551.2014.09.1615

Zhao, Z., Gu, H., Sheng, X., Yu, H., Wang, J., Zhao, J., et al. (2014b). Genetic diversity and relationships among loose-curd cauliflower and related varieties as revealed by microsatellite markers. *Sci. Hortic.* 166, 105–110. doi: 10.1016/j.scienta.2013.12.024

Zhao, Z., Gu, H., Wang, J., Sheng, X., and Yu, H. (2013). Development and comparison of quantitative methods to evaluate the curd solidity of cauliflower. *J. Food Eng.* 119, 477–482. doi: 10.1016/j.jfoodeng.2013.06.025

Zhou, X., Xia, Y., Ren, X., Chen, Y., Huang, L., Huang, S., et al. (2014). Construction of a SNP-based genetic linkage map in cultivated peanut based on large scale marker development using next-generation double-digest restriction-site-associated DNA sequencing (ddRADseq). *BMC Genomics* 15:351. doi: 10.1186/1471-2164-15-351

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhao, Gu, Sheng, Yu, Wang, Huang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 30 August 2016
doi: 10.3389/fpls.2016.01258

Genetic and Epigenetic Alterations of *Brassica nigra* Introgression Lines from Somatic Hybridization: A Resource for Cauliflower Improvement

*Gui-xiang Wang[1], Jing Lv[1,2,3], Jie Zhang[1], Shuo Han[1], Mei Zong[1], Ning Guo[1], Xing-ying Zeng[1], Yue-yun Zhang[1], You-ping Wang[2] and Fan Liu[1]**

[1] Beijing Vegetable Research Center, Beijing Academy of Agriculture and Forestry Sciences, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (North China), Ministry of Agriculture, Beijing, China, [2] Yangzhou University, Yangzhou, China, [3] Zhalute No.1 High School, Tongliao, China

Broad phenotypic variations were obtained previously in derivatives from the asymmetric somatic hybridization of cauliflower "Korso" (*Brassica oleracea* var. *botrytis*, 2*n* = 18, CC genome) and black mustard "G1/1" (*Brassica nigra*, 2*n* = 16, BB genome). However, the mechanisms underlying these variations were unknown. In this study, 28 putative introgression lines (ILs) were pre-selected according to a series of morphological (leaf shape and color, plant height and branching, curd features, and flower traits) and physiological (black rot/club root resistance) characters. Multi-color fluorescence *in situ* hybridization revealed that these plants contained 18 chromosomes derived from "Korso." Molecular marker (65 simple sequence repeats and 77 amplified fragment length polymorphisms) analysis identified the presence of "G1/1" DNA segments (average 7.5%). Additionally, DNA profiling revealed many genetic and epigenetic differences among the ILs, including sequence alterations, deletions, and variation in patterns of cytosine methylation. The frequency of fragments lost (5.1%) was higher than presence of novel bands (1.4%), and the presence of fragments specific to *Brassica carinata* (BBCC 2*n* = 34) were common (average 15.5%). Methylation-sensitive amplified polymorphism analysis indicated that methylation changes were common and that hypermethylation (12.4%) was more frequent than hypomethylation (4.8%). Our results suggested that asymmetric somatic hybridization and alien DNA introgression induced genetic and epigenetic alterations. Thus, these ILs represent an important, novel germplasm resource for cauliflower improvement that can be mined for diverse traits of interest to breeders and researchers.

Keywords: introgression lines, genetic diversity, *Brassica oleracea*, *Brassica nigra*, somatic hybridization, epigenetic variation, cauliflower

OPEN ACCESS

Edited by:
Narendra Tuteja, International Centre for Genetic Engineering and Biotechnology, India

Reviewed by:
Anca Macovei, University of Pavia, Italy
Carl Gunnar Fossdal, Norwegian Institute of Bioeconomy Research, Norway

****Correspondence:***
Fan Liu
liufan@nercv.org

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *13 May 2016*
Accepted: *08 August 2016*
Published: *30 August 2016*

Citation:
Wang G-x, Lv J, Zhang J, Han S, Zong M, Guo N, Zeng X-y, Zhang Y-y, Wang Y-p and Liu F (2016) Genetic and Epigenetic Alterations of Brassica nigra Introgression Lines from Somatic Hybridization: A Resource for Cauliflower Improvement. Front. Plant Sci. 7:1258. doi: 10.3389/fpls.2016.01258

INTRODUCTION

Cauliflower (*Brassica oleracea* var. *botrytis*, 2*n* = 18, CC genome) is a major vegetable crop valued worldwide for its nutrition and flavor. As with other modern crop species, intense selection for preferred traits, and a tendency for inbreeding have resulted in low genetic diversity among cauliflower breeding resources. To address this problem, related species such as black mustard

(*Brassica nigra*, $2n = 16$, BB genome), with a large reservoir of genes conferring desirable characteristics, have been proposed as a valuable source of genetic diversity for *Brassica* crop improvement. In principle, this diversity can be transferred into crops via sexual hybridization and subsequent backcrossing. However, in practice, genetic manipulation using distantly related plants has been severely restricted by difficulties in creating the initial sexual hybrid and by sterility issues in the early backcross generations (Glimelius, 1999; Jauhar, 2006). Asymmetric somatic hybridization is a potential alternative for gene transfer from relatives to cultivated crops, especially when wide crosses are not applicable (Gerdemann et al., 1994; Xia, 2009). The Brassicaceae is a model plant family commonly used for somatic hybridization (Glimelius, 1999; Navrátilová, 2004), resulting in interspecies or even intertribal hybrids. Moreover, agronomically important traits, such as disease resistance and specific fatty acid compositions, have been successfully integrated into the crops (Gerdemann et al., 1994; Hansen and Earle, 1994, 1995, 1997; Wang et al., 2003; Tu et al., 2008; Scholze et al., 2010).

Brassica is also an excellent model for the study of allotetraploid speciation. *Brassica* is closely related to the classic plant model *Arabidopsis* and a wealth of germplasm exists from species in this genus. Three diploid species are widely cultivated: *Brassica rapa* (AA, $x = 10$), *B. nigra* (BB, $x = 8$), and *B. oleracea* (CC, $x = 9$). Recent interspecific hybridization among these species has resulted in polyploidization and the production of three allotetraploid species: *Brassica napus* (AACC, $x = 19$), *Brassica juncea* (AABB, $x = 18$), and *Brassica carinata* (BBCC, $x = 17$) (Nagaharu, 1935).

The formation of allotetraploids is an influential mode of speciation in flowering plants, often accompanied by rapid, extensive genomic and epigenetic changes that globally alter gene expression, termed "genomic shock." These include fragment gain and loss through chromosome rearrangement or the activation of transposable elements (Kashkush et al., 2003; Kraitshtein et al., 2010), extensive alteration of DNA cytosine methylation (Zhao et al., 2011; Gautam et al., 2016), histone modification (Ha et al., 2011), and changes to small RNA (Ha et al., 2009). However, genomic-shock-induced changes during sexual polyploid synthesis should be distinct from the changes that occur from genomic shock during somatic hybridization. Somatic hybrids combine both the nuclear and cytoplasmic genomes within a single cell. Therefore, the introgression of chromatin segments via asymmetric somatic hybridization likely occurs via non-homologous end-joining of fragmented genome pieces, rather than by the homologous recombination that occurs through sexual reproduction (Liu et al., 2015). Moreover, the epigenetic states of somatic cells and gametal cells tend to differ, given that gametal cells are more conserved to ensure genetic fidelity (Bird, 1997, 2002; Liu et al., 2015). Thus, the variations induced by "somatic genomic shock" likely have unique genetic and epigenetic characteristics compared with "sexual genomic shock." However, little data exist on the exact nature of these differences, or if the types of genomic shock indeed differ in their effects.

A number of hybrid progenies have been regenerated from asymmetric somatic hybrids between cauliflower "Korso" and black mustard "G1/1." The hetero-cytoplasmic nature of these hybrids was confirmed through the finding that most progenies have chloroplast genomic components from "G1/1" and mitochondrial DNA from "Korso," while mitochondrial genome recombination occurred in a few hybrids (Wang G. X. et al., 2011). Among the progenies, dozens of lines containing 18 chromosomes showed some obvious characters of "G1/1" origin, such as waxless leaves with lobes and ears, as well as green petioles; however, they were also densely covered with trichomes and exhibited purple stems and leaf veins, characteristic of "Korso." This combination of characters indicated that they were putative introgression lines (ILs). Although hybrid synthesis primarily focused on the transfer of disease-resistance genes from "G1/1" into cultivated cauliflower (Wang G. X. et al., 2011) because the former is resistant to several *Brassica* pathogens (Westman et al., 1999), the putative ILs contained considerable genetic diversity for many traits beyond disease resistance. Therefore, these lines should possess many characters that differ from near-isogenic lines and feature discrete portions of "G1/1" chromatin in a "Korso" genetic background. Additionally, asymmetric somatic introgression should induce further diverse genomic variations involving change in DNA sequences and epigenetic modifications. To verify this, we used 12 putative ILs shown to be phenotypically stable for several generations to characterize genetic and epigenetic alterations from "somatic genomic shock."

MATERIALS AND METHODS

Plant Materials

Asymmetric somatic hybridization between cauliflower "Korso" and black mustard "G1/1" resulted in the establishment of 117 individual hybrids in the field; of these, 13 fertile plants with preferred traits were selected for continued selfing and backcrossing (Wang G. X. et al., 2011). Hundreds of derivatives were obtained following year-by-year selection (since 2006) that combined phenotypic observation, cytological and molecular analysis, as well as pathogen-resistance assays (**Figure 1**). In this study, 28 putative ILs were chosen for phenotypic diversity and chromosomal constitution analysis, based on possession of cauliflower-like morphology and phenotypic stability over three generations. Next, 12 putative ILs (IL1-12) were used for genetic an epigenetic analysis. Chinese cabbage "Asko" (*B. rapa*, AA, $2n = 20$) and tetraploid *B. carinata* (BBCC, $2n = 34$) were chosen as reference materials for molecular and chromosomal analysis. The selfed seeds from ILs, parents, and *B. carinata* were sown in a greenhouse and are available upon request.

Multi-Color FISH Analysis

Multi-color FISH was conducted on meiosis metaphase I chromosomes of parent lines and putative ILs to confirm their chromosomal composition. *B. nigra* and CentBr (centromeric-specific tandem repeats of *Brassica*; accession numbers: CW978699 and CW978837, respectively; Wang G. et al., 2011) DNA were selected as probes and labeled with biotin (dig)-14-dUTP (Roche, Indianapolis, IN) using nick translation with an average length of 500 bp. Slides were prepared following a previously published protocol (Zhong et al., 1996) with minor

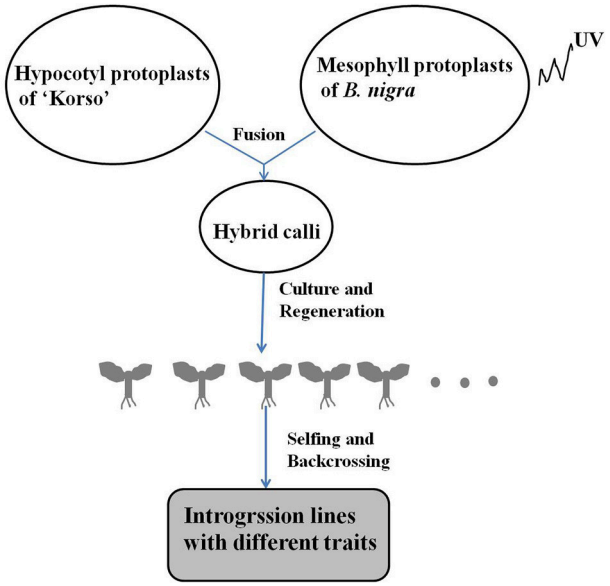


FIGURE 1 | Generation of introgression lines via somatic hybridization.

modifications. To decompose the cell walls of pollen mother cells, anthers approximately 1–3 mm in length were digested at 37°C for 3 h in an enzyme mixture containing 6% cellulose R-10 and 6.5% pectinase (Sigma, solution in 40% glycerol). Previously described methods (Wang G. et al., 2011) were followed for FISH analysis. Hybridized probes were visualized using an FITC-conjugated avidin antibody or rhodamine-conjugated anti-digoxin antibody (Roche, Indianapolis, IN). Chromosomes were counterstained using 0.1 mg/mL DAPI (Vector Laboratories, Burlingame, CA). Images of the signals and chromosomes were captured using a CCD camera (QIMAGING, RETIGA-SRV, FAST1394) attached to a Nikon Eclipse 80i epifluorescence microscope (Tokyo, Japan). Image contrast and brightness were adjusted in Adobe Photoshop (8.0).

SSR, AFLP, and MSAP Fingerprinting

Simple sequences repeat (SSR), amplified fragment length polymorphism (AFLP), and methylation-sensitive amplification polymorphism (MSAP) fingerprinting were performed following previously described methods (Shaked et al., 2001; Liu et al., 2015). Sixty-five SSR, 77 AFLP, and 16 MSAP primer combinations were used; for SSR and AFLP, markers were collected from the A genome of *B. rapa* due to the limited number of cytological and genetic studies on the B genome of *B. nigra* (Tables S1–S3). For AFLP and MSAP, two independent technical replicates were performed for each of the three biological replicates. Only clear and reproducible bands larger than 100 bp were scored.

RESULTS

Phenotypic Diversity of Introgression Lines

The 28 ILs were extremely diverse in multiple morphological and physiological traits (**Table 1**), including distinct variations in leaf or flower morphology, fertility, and resistance to pathogens (i.e., black rot or clubroot). Here we describe diagnostic differences in flower heads as an example. Four phenotypic groups were apparent based on flower curd characters (**Figure 2**). In the first group (type I), flower heads were white, compact, and hemispheric, with a wheel-like radial arrangement; flower buds were small, granular, and close. The second group (type II) possessed flat, loose flower heads with no obvious wheel-like arrangement; flower buds were light yellow, tiny, and soft, while the pedicel was green and slightly longer than Type I pedicels. The third group (type III) possessed relatively flat flower heads and pale yellow-green color; flower buds were small, loose, soft, and distributed in clusters, but with no obvious ball flower formation. The pedicel was green and long. The flower heads of the fourth group (type IV) were relatively flat, yellow green, with small flower buds that were fine and soft.

Chromosomal Constitution of the Putative Introgression Lines

Figure 3 shows the results of multi-color FISH on chromosomes of "G1/1," "Korso," and 28 putative ILs. The two parent lines exhibited different hybridization patterns with both probes. As expected, when using "G1/1" genomic DNA, all 16 "G1/1" chromosome pairs exhibited red signals (**Figure 3A**). In "Korso," red signals (2 strong and 2 weak; **Figure 3B**) were detected on only two pairs of satellite chromosomes that likely correspond to the hybridization signals of 45SrDNA (Fukui et al., 1998).

No signals were detected on "G1/1" chromosomes using the CentBr probe. In contrast, seven pairs of strong signals and one pair of weak signals were detected on "Korso" chromosomes (**Figure 3B**); the latter was hybridized to chromosomes that also exhibited strong red signals. Only 1 "Korso" chromosome pair exhibited no signals (**Figure 3B**).

Most cells in the putative ILs contained 18 chromosomes and exhibited hybridization patterns identical to "Korso." The results of genomic *in situ* hybridization (GISH) found no "G1/1" chromosomal segments present in these putative ILs, likely because GISH cannot detect introgressions from *Brassica* chromosomes due to their small size and compact structure.

Genetic Sequence Variation Analyzed Using SSR and AFLP Markers

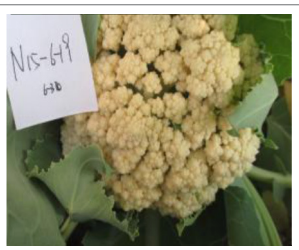
We analyzed the sequence introgression and variation in 12 ILs (IL1–12) using 65 SSR and 77 AFLP markers that were polymorphic between the parents. The results of SSR and AFLP profiling revealed 1799 loci from both parents, including 682 (37.9%) those were polymorphic. The IL profiles were similar to the "Korso" profile, but loss of 5–13 "Korso"-specific bands were common. Additionally, all putative ILs were confirmed to possess anywhere from 3 to 33 "G1/1"-derived fragments. However, we noted that in 3 ILs, SSR markers failed to detect "G1/1"-derived bands but AFLP profiling revealed 20–25 "G1/1"-specific loci, indicating that the latter method is more effective when the introgression size is small. Next, all 12 ILs contained *B. carinata*-specific loci (35–49), and we were able to amplify 1–7 new bands from them.

TABLE 1 | Phenotypic traits associated with introgression lines (IL1-12) and their parents.

Materials	Leaf traits	Curd traits	Other special traits	Curd morphology
"Korso"	Ovate; shallowly serrated margins; lacking auricles; waxy, glossy, and hairless surface	White and firm curd; short pedicel	Susceptible to both black rot *Xcc*15 and clubroot race 4; potassium content: 156 mg/100 g; phosphorus content: 36 mg/100 g	
"G1/1"	Almost triangular; possesses auricles; green; and wax-free surface	No curd; direct flowering	Resistant to both *Xcc*15 and clubroot race 4	
IL1	Grayish green; possesses auricles	Yellow and firm curd; short pedicel		
IL2	Pale grayish green; shallowly serrated margins; possesses auricles	White and firm curd; purple florets, short pedicel	Resistant to *Xcc*15	
IL3	Green; serrated margins; no auricles	Pale and greenish curd; long pedicel	Early inflorescent; phosphorus content: 73 mg/100 g	
IL4	Light grayish green; serrated margins; no auricles	White and firm curd; short pedicel		

(Continued)

TABLE 1 | Continued

Materials	Leaf traits	Curd traits	Other special traits	Curd morphology
IL5	light green; wavy margin; possesses auricles	Pale green and soft curd; long pedicel		
IL6	Pale grayish green; wavy margin; wing-like	White; firm, medium curd; white florets	Resistant to clubroot race 4	
IL7	Pale grayish green; elongated; no auricles	No obvious curd; flower buds in clusters with leaf folder		
IL8	Light grayish green; serrated margins; no auricles	Pale yellow and scattered curd; long pedicel		
IL9	Green; serrated margins; possesses auricles	Yellowish and loose curd; long pedicel	Potassium content: 236 mg/100 g	
IL10	Pale grayish green; serrated margins; possesses auricles	Creamy white and loose curd; purple flower buds; long pedicel		

(Continued)

TABLE 1 | Continued

Materials	Leaf traits	Curd traits	Other special traits	Curd morphology
IL11	Similar to parent "Korso"	Creamy white and soft tomentose curd; short pedicel	Resistant to *Xcc*15	
IL12	Pale grayish green; serrated margins; possesses auricles	Yellowish green and soft curd; purple flower buds; long pedicel	Early inflorescent	

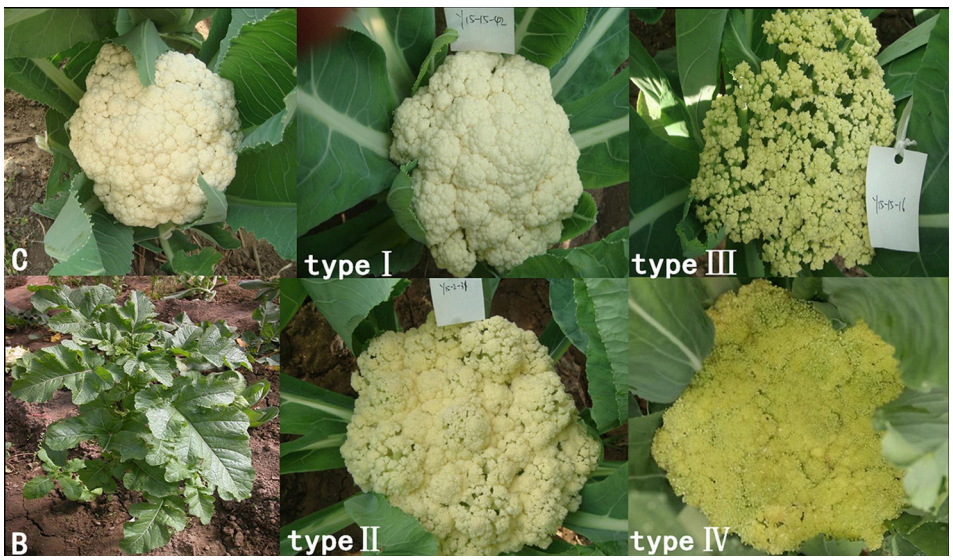


FIGURE 2 | Four types (I–IV) of representative flower heads in introgression lines. C: "Korso," B: "G1/1."

These results clearly showed that all 12 lines were ILs and that genomic-sequence variations were common among them. Notably, IL genomes were not simply similar to the "Korso" genome but with some "G1/1"-derived fragments added. For example, we noted that some ILs lacked bands present in "Korso," or possessed fragments lacking in "Korso." Moreover, several *B. carinata*-specific bands were present in all ILs, hinting at the occurrence of common changes in the early stages of both somatic hybridization and naturally occurring hybridization.

DNA profiling of the ILs revealed 1493 "Korso," 1423 "G1/1," and 1559 *B. carinata* fragments; 1117 of these were shared by both parents (**Figure 4**). The number of fragments present in the profiles of the 12 representative ILs ranged from 3 to 33, with an overall frequency of 7.5%, whereas the number of fragments lost ranged from 7 to 32, with an overall frequency of 5.1% (**Table 2**). Note that the percentages of "G1/1"-fragment presence and "Korso"-fragment loss were respectively based on the 306 "G1/1" fragments that did not co-migrate with "Korso" and the 376 "Korso" fragments that did not co-migrate with "G1/1." In addition, the 12 ILs possessed 1–7 new fragments (1.4%) that were not present in the "Korso" profile, and 4–49 (15.5%) *B. carinata*-specific fragments from the 226 that did not co-migrate with "Korso" and "G1/1."

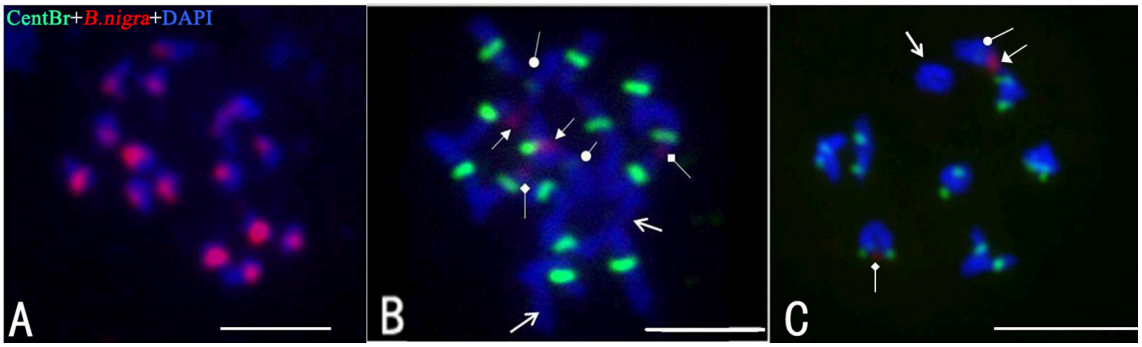


FIGURE 3 | Fluorescence *in situ* hybridization (FISH) analysis of the introgression lines using CentBr and "G1/1" genomic DNA as probes. (A) Meiosis metaphase I of "G1/1." **(B)** Mitosis metaphase I of "Korso." **(C)** Meiosis prometaphase I of IL2. Triangle- and diamond-headed pointers indicate strong and weak red signals, respectively; circle-headed pointers indicate weak green signals; regular arrows indicate chromosomes with no signals. Bar = 5 μm.

Epigenetic Changes in Methylation Patterns Analyzed with MSAP

The MSAP analysis used isoschizomers *Msp*I and *Hpa*II, which differ in their sensitivity to cytosine methylation (Shaked et al., 2001; Liu et al., 2015). Our comparison of the profiles generated from *Eco*RI-*Msp*I and *Eco*RI-*Hpa*II digestion revealed three kinds of band classifications (**Figure 5**). First, A type was characterized by fragments from both the *Eco*RI-*Msp*I and *Eco*RI-*Hpa*II digestions, representing non-methylated sites (or methylation within a single strand). Second, B type was characterized by half-methylation sites, present in *Eco*RI-*Hpa*II digestion but not in *Eco*RI-*Msp*I digestion. Third, C-type bands stemmed from complete methylation sites present in *Eco*RI-*Msp*I digestion but not in *Eco*RI-*Hpa*II digestion. Using 16 MSAP primer combinations amplifying clear and stable bands, we detected 401 fragments in parents, with 290 present in "Korso" but not in "G1/1."

Table 3 summarizes variation in methylation patterns and their heredity in parents and the ILs. The average frequency of hypermethylated loci in different ILs was 12.4%, while the average frequency of hypomethylated loci was 4.8%.

DISCUSSION

Somatic Hybridization and Introgression Can Potentially Increase Genetic Diversity

The low genetic diversity in modern Brassica varieties is concerning because it reduces potential genetic gains in breeding programs. To combat this problem, ILs containing fragments from related species can be used to generate improved cultivars. This method has proven successful in many crops, including wheat (Liu et al., 2007), rice (Rangel et al., 2008), potato (Chavez et al., 1988), eggplant (Mennella et al., 2010), *B. napus* (Primard-Brisset et al., 2005; Leflon et al., 2007), barley (Johnston et al., 2009), tomato (Menda et al., 2014), and rye grass (Roderick et al., 2003).

In this study, we generated a set of ILs using somatic hybridization between cauliflower and black mustard. The original objective was to select disease- and drought-resistant ILs for use in a Brassica breeding program. In fact, these ILs showed considerable genetic diversity in many morphological and physiological traits beyond resistance. Some of these traits are useful for increasing cauliflower adaptation to stressors. For instance, IL2 and IL6 respectively exhibited strong resistance to black rot and clubroot, major Brassica diseases with worldwide distribution. Other ILs exhibited commercially desirable traits such as early maturation and high nutritional quality, especially high phosphorus and potassium content (IL3 and IL9), as well as less desirable traits such as loose flower heads, Early inflorescent, and long pedicel. This diversity indicates that all of the ILs could be a source of new alleles for developing Brassica cultivars. Thus, somatic hybridization provides an effective and efficient means of achieving introgression into a crop species from its wild relatives and should continue to be important in producing novel germplasm for breeding programs.

Combining Cytological and Molecular Methods Effectively Identified Brassica Introgression Lines

Although GISH has been successfully used to investigate genomic relationships in several agriculturally important genera, including Brassica (e.g., Lysak et al., 2005), and has detected introgressions in crops such as wheat (e.g., Liu et al., 2015), GISH was unable to detect black mustard chromosomal segments in our ILs. We believe this outcome was likely due to the small and compact Brassica chromosomes or else very small introgression sizes. We dealt with this issue by first using multi-color FISH to identify chromosome composition and then combining SSR and AFLP markers to detect "G1/1" sequence introgression. Primer sources and amplification results from the genomes of *B. rapa*, *B. nigra*, *B. oleracea*, and *B. carinata* are shown in Table S1. As those data indicate, the primers are usable in multiple Brassica species and can detect abundant polymorphism. Moreover, compared with the SSR primers, AFLP markers were more effective when working with small introgressions, such as those in our ILs (**Table 2**). In sum, our study demonstrated that Brassica ILs can be clearly identified with a combination of cytological and molecular methods.

TABLE 2 | DNA profiles and phenotypic summary of partial introgression lines (IL1-12).

Introgression lines (ILs)	No. of *B. nigra*-specific bands and markers	Deletions in ILs	New bands in ILs	*B. carinata*-specific bands and markers	Phenotype
IL1	25: (23 AFLP), BrID10227, nia-m037a	12: (8 AFLP), BrID10205, nia-m037a, At3g55005a, cnu-m472a	7: (6 AFLP), BrID10223	40: (32 AFLP), BrID10205, BrID101161, BrID10223, At3g55005a, cnu-m472a, cnu-m052a	Leaf color and leaf type like the cauliflower; small, white, and firm flower head
IL2	26: (25 AFLP), BrID10227	13: (8 AFLP), BrID10205, BrID10381, BrID10173, nia-m037a	6 AFLP	35: (32 AFLP), BrID10985, BrID101161, BrID10381	Leaf color and leaf type like the cauliflower; small, white, and firm flower head
IL3	31: (29 AFLP), BrFLC2(R2), BrID10227	16: (9 AFLP), BrID10277, BrID10205, BrFLC2(R2), nia-m037a, cnu-m472a, nia-m063	6 AFLP	49: (41 AFLP), BrID10277, BrID10205, BrFLC2(R2), BrID101161, cnu-m472a, cnu-m052a	Leaf color and leaf type like the cauliflower; small, white, and firm flower head
IL4	20: (18 AFLP), BrFLC2(R2), BrID101081	23: (8 AFLP), BrID10277, BrID10205, BrFLC2(R2), BrID10881, BrID101157, BrID101183, BrID90311, BrID10493, BrID10223, BrID10173, nia-m037a, nia-m063a	6 AFLP	40: (32 AFLP), BrID10277, BrFLC2(R2), BrID90430, BrID101161, cnu-m052a	Leaf color and leaf type like the cauliflower; small, white, and firm flower head
IL5	22: (21 AFLP), At4g36960	26: (9 AFLP), BrID10277, BrID10205, BrID10881, BrID101157, BrID101161, BrID10971, BrID10381, BrID90311, BrID10493, BrID10173, nia-m037a, At3g55005a, nia-m063a, At3g51260	6 AFLP	40: (33 AFLP), BrID90430, BrID10381, BrID10493, At3g55005a, nia-m063a, cnu-m052a	Leaf color and leaf type like the cauliflower; loose, light yellow flower head; some light purple areas
IL6	27: (25 AFLP), BrID10227, At4g36960	11: (7 AFLP), BrID10205, BrID10493, cnu-m472a	6 AFLP	35: (31 AFLP), BrID10493, cnu-m472a, cnu-m052a	Leaf color and leaf type like the cauliflower; small, white, and firm flower head
IL7	32: (31 AFLP), At4g36960	30: (9 AFLP), BrID10277, BrID10205, BrID10881, BrID10381, BrID101183, BrID10971, BrID90311, BrID10103, BrID10493, BrID10101, BrID10275, BrID101199, nia-m037a, At3g55005a, cnu-m472a, At3g51260	6 AFLP	48: (40 AFLP), BrID10277, BrID90430, BrID101161, BrID10381, At3g55005a, cnu-m472a, cnu-m052a	Leaf color and leaf type like the cauliflower; white and firm flower head
IL8	27: (25 AFLP), BrID10985, BrFLC2(R2)	32: (13 AFLP), BrID10205, BrFLC2(R2), BrID101157, BrID101183, BrID90311, BrID10493, BrID10101, BrID10971, BrID10275, BrID101199, BrID10173, At1g58220, nia-m037a, At3g55005a, At4g36960, At3g51260	6 AFLP	49: (41 AFLP), BrID10277, BrID10985, BrID10205, BrFLC2(R2), BrID90430, At3g55005a, cnu-m052a	Waxy leaf and loose flower head
IL9	33: (31 AFLP), BrFLC2(R2), BrID10227	17: (11 AFLP), BrID10277, BrID10205, BrFLC2(R2), BrID101157, nia-m037a, cnu-m472a	6 AFLP	45: (38 AFLP), BrID10277, BrID10205, BrFLC2(R2), BrID101161, cnu-m472a, cnu-m052a	Loose flower head

(Continued)

TABLE 2 | Continued

Introgression lines (ILs)	No. of *B. nigra*-specific bands and markers	Deletions in ILs	New bands in ILs	*B. carinata*-specific bands and markers	Phenotype
IL10	25 AFLP	17: (8 AFLP), BrID10881, BrID101157, BrID101183, BrID10971, BrID10493, nia-m037a, At3g55005a, At3g51260	6 AFLP	34: (31 AFLP), BrID101161, At3g55005a, At3g51260	Waxy leaf and loose flower head
IL11	3: BrID10227, nia-m037a, At3g55005a	7: BrID10277, BrID10205, BrID101157, BrID10381, nia-m037a, At3g55005a, Ni4-B10	1: Ni4-B10	4: BrID10277, BrID10205, BrID10381	Ligule present; small, tight flower head
IL12	5: BrID10277, BrID10985, nia-m037a, At3g55005a, At2g38130	27: BrID10205, BrFLC2(R2), BrID10881, BrID101157, BrID101183, BrID90311, BrID10971, BrID10103, BrID10493, BrID10101, BrID10275, BrID101199, BrID10173, At1g58220, nia-m037a, At3g55005a, cnu-m472a, nia-m063a, Ni4-B10, At3g51260	2: BrID10223, Ni4-B10	5: BrID10277, BrID10223, BrID90430, At2g38130, nia-m063a	Gray-green leaf color; firm, small flower head
Average	23.0	19.3	5.3	35.3	
Frequency	7.5%	5.1%	1.4%	15.5%	

Genetic Changes from Somatic Hybridization in the Brassica Introgression Lines

The extensive phenotypic variation of the ILs was caused by genetic and epigenetic changes in the ILs relative to their parents. However, black mustard introgression was actually more limited (7.5%; **Table 2**) than other ILs, such as the somatic hybrids between wheat and tall wheatgrass (Liu et al., 2015). We believe the low introgression was probably due to suboptimal UV treatment. Thus, treatment factors such as irradiation time and intensity will require further testing and validation to achieve higher introgression rates. At the same time, the amount of introgression observed may be enough to achieve considerable genomic diversity. For example, in stable rice introgression lines derived from intergeneric hybridization (followed by successive selfing) between rice (*Oryza sativa* L.) and a wild relative (*Zizania latifolia* Griseb.), very low introgression (0.1%) by foreign DNA was enough to trigger up to 30% of the genomic changes (Wang et al., 2005). The mechanism underlying these changes is unclear, but some researchers have suggested that they happen very early, through a cryptic pathway that differs from conventional or unorthodox meiotic recombination of homeoalleles between rice and Zizania (Zhang et al., 2013).

Similar to previous findings (Wang et al., 2005), we also observed sequence loss and new bands (**Table 2**) following somatic hybridization, backcrossing, and self-pollination (**Figure 1**). The frequency of fragments lost (5.1%) was higher than presence of novel bands (1.4%). It is not entirely clear why the percentage of novel bands was so much lower, but one reason may be that some of them were *B. carinata*-specific, and *B. carinata*-specific fragments occurred at a relatively high frequency (15.5%). Thus, future investigations may need to confirm the exact nature of the lost and novel bands, as well as exclude any *B. carinata*-specific loci.

TABLE 3 | Variation in cytosine methylation across the partial introgression lines (IL1-12).

Introgression lines (ILs)	Hypermethylation in ILs	Frequency (%)	Hypomethylation in ILs	Frequency (%)
IL1	42	14.5	7	2.4
IL2	36	12.4	9	3.1
IL3	35	12.1	8	2.8
IL4	39	13.5	22	7.6
IL5	28	9.7	18	6.2
IL6	32	11.0	17	5.9
IL7	38	13.1	12	4.1
IL8	28	9.7	10	3.5
IL9	34	11.7	8	2.8
IL10	42	14.5	22	7.6
IL11	40	13.8	20	6.9
IL12	36	12.4	13	4.5
Average	35.8	12.4	13.8	4.8

Our results were consistent with previous reports in artificial hybrid *B. carinata* (BBCC, $x = 17$), produced via crossing and polyploidization between *B. nigra* (BB, $x = 8$) and *B. oleracea* (CC, $x = 9$). Around 47% of the hybrid *B. carinata* (BBCC) genome possessed isoenzyme sites specific to natural allotetraploid *B. carinata* (Jourdan and Salazar, 1993). Given

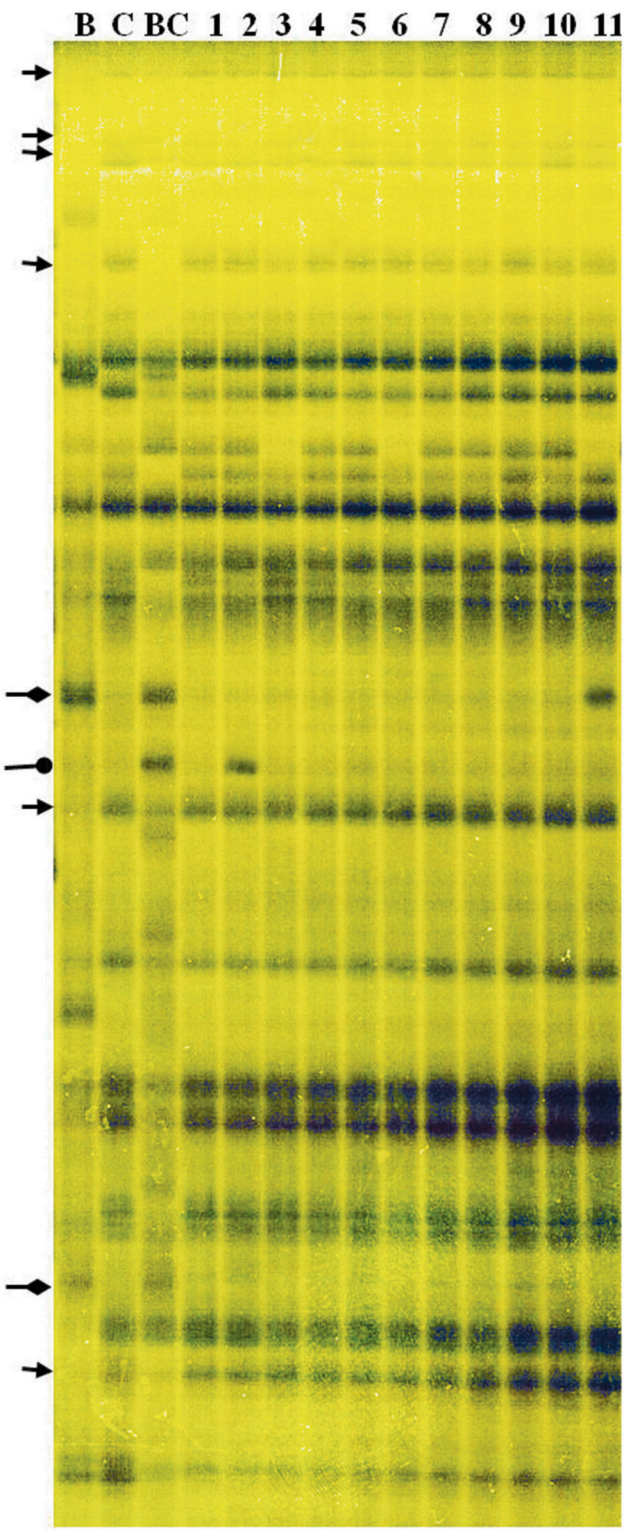


FIGURE 4 | Amplified fragment length polymorphism (AFLP) fingerprinting of the parents and ILs, with primer pair E45–M57. C: "Korso," B: "G1/1," BC: *B. carinata*, 1–11: introgression lines. Arrows indicate "Korso"-specific bands, diamond-headed pointers indicate "G1/1"-specific bands, and circle-headed pointers indicate *B. carinata*-specific bands.

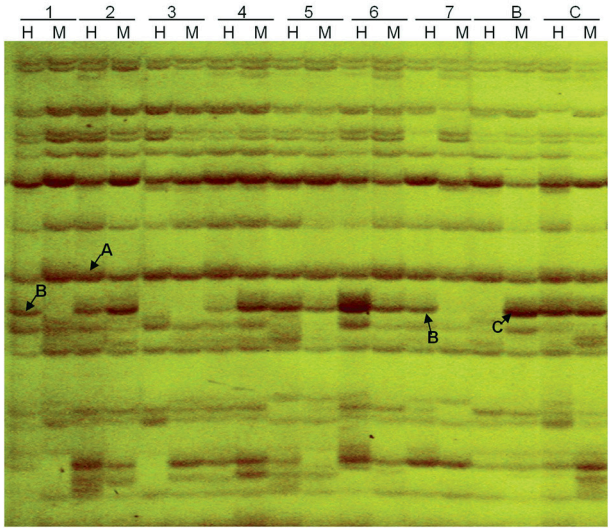


FIGURE 5 | Methylation-sensitive amplified polymorphism (MSAP) profiling based on primer combinations ETTC + HMTCC. 1–7: introgression lines, B: "G1/1," C: "Korso," H: DNA cleaved with EcoRI + HpaII, M: DNA cleaved with EcoRI + MspI. A, B, and C indicate three separate band types.

this similarity, we propose that the early generations of artificial synthesis and natural polyploid evolution of both genomes may have experienced the same change events. Supporting this idea, previous analyses of recombination in other genera, such as the homoploid hybrid sunflower (*Helianthus anomalus*), showed that newly synthesized hybrids converged on the linkage pattern of wild hybrids within five generations (Rieseberg et al., 1996).

Based on our results and those of previous studies, we propose that genetic changes in Brassica ILs occur via the following mechanisms. First, alien genetic elements are incorporated into cauliflower early in hybridization, through the activation of transposable elements or DNA methylation, as suggested in rice introgression lines (Wang et al., 2005). Next, chromosomal rearrangements occur through intergenomic translocations or transpositions (Udall et al., 2005) and homoeologous pairing (Leflon et al., 2006; Nicolas et al., 2007). During Brassica evolution, ancestral group rearrangements contributed to the high homology found especially in the three linkage groups of the B genome (B5, B6, and B4) and the A genome (A5, A6, and A4) (Panjabi et al., 2008). Finally, consecutive selection results in rapid sequence elimination, as reported in synthetic hybrids and hybrids less related to the original parents (Osborn et al., 2003; Pires et al., 2004; Gaeta et al., 2007; Zhang et al., 2013).

Changes to DNA Methylation in the Brassica Introgression Lines

Cytosine methylation plays an important role in epigenetic gene regulation at both the transcriptional and the post-transcriptional levels (Paszkowski and Whitham, 2001). Our ILs exhibited fairly high proportions of change to methylation patterns (17.2%; **Table 3**), compared with several other plant

hybrid/allopolyploid systems analyzed to date using MSAP. For instance, in resynthesized allotetraploid Arabidopsis suecica, 8.3% of the fragments experienced methylation changes (Madlung et al., 2002), whereas in newly synthesized allohexaploid wheat, ∼13% of the loci saw alterations to cytosine methylation (Shaked et al., 2001). In contrast, a 23.6% change to cytosine methylation was observed in asymmetric somatic wheat introgressions (Liu et al., 2015). Taken together, these data allow us to conclude that somatic hybridization induced a broad spectrum of cytosine-methylation changes that perturbed gene expression to a larger extent than allopolyploidization.

Consistent with previous findings (Zhang et al., 2013), changes to hypermethylation (9.7–13.5%; **Table 3**) occurred more frequently than changes to hypomethylation (2.4–7.6%). Zhang et al. (2013) proposed that this pattern may be caused by the activation and subsequent silencing of some transposable elements after hybridization, but this hypothesis requires more data for confirmation. Additionally, the methylation patterns in ILs may be related to changes in enzymatic machinery, as well as the expression of siRNAs and long noncoding RNAs (Goll and Bestor, 2005; Wierzbicki et al., 2008; Zhang et al., 2013).

In conclusion, somatic hybridization mimics many of the genetic alterations induced by polyploidization or sexual wide hybridization, but to a stronger extent and with considerably less time. Therefore, somatic hybridization is both effective and efficient in achieving introgression of a crop species and its wild relatives. Moreover, this approach provides a potential means to explore the genetic and epigenetic events induced by "somatic genomic shock" (Liu et al., 2015).

AUTHOR CONTRIBUTIONS

GW conceived the research work and wrote the paper. FL provided plant materials and intellectual advice on the project. JL performed molecular and FISH experiments. JZ and SH performed the field management and character survey. XZ performed the methylation-sensitive amplified polymorphism analysis. NG participated in the technical guidance of relative experiment. MZ and YZ performed materials classification and preservation. YW provided technical guidance for pathogen-resistance assays.

ACKNOWLEDGMENTS

This work was supported by the Natural Science Foundation of China (No. 31000538), the Natural Science Foundation of Beijing, China (No. 6142012), the Project of Technology Innovation Ability from Beijing Academy of Agriculture and Forestry Sciences (KJCX20140111), and the Youth Science Research Foundation of Beijing Academy of Agriculture and Forestry Sciences (No. QNJJ201601). National Key Research and Development Project (2016YFD010 0204-14).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01258

REFERENCES

Bird, A. P. (1997). Does DNA methylation control transposition of selfish elements in the germline. *Trends Genet.* 13, 469–470. doi: 10.1016/S0168-9525(97)01310-3

Bird, A. P. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21. doi: 10.1101/gad.947102

Chavez, R., Brown, C. R., and Iwanaga, M. (1988). Application of interspecific sesquiploidy to introgression of PLRV resistance from non-tuber-bearing *Solanum etuberosum* to cultivated potato germplasm. *Theor. Appl. Genet.* 76, 497–500. doi: 10.1007/bf00260898

Fukui, K., Nakayama, S., Ohmido, N., Yoshiaki, H., and Yamabe, M. (1998). Quantitative karyotyping of three diploid Brassica species by imaging methods and localization of 45S rDNA loci on the identified chromosomes. *Theor. Appl. Genet.* 96, 325–330. doi: 10.1007/s001220050744

Gaeta, R. T., Pires, J. C., Iniguez-Luy, F., and Osborn, L. E. (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* 19, 3403–3417. doi: 10.1105/tpc.107.054346

Gautam, M., Dang, Y., Ge, X., Shao, Y., and Li, Z. (2016). Genetic and epigenetic changes in oilseed rape (*Brassica napus* L.) extracted from intergeneric allopolyploid and additions with *Orychophragmus*. *Front. Plant Sci.* 7:438. doi: 10.3389/fpls.2016.00438

Gerdemann, M. K., Sacristán, M. D., Braatz, C., and Schieder, O. (1994). Utilization of asymmetric somatic hybridization for the transfer of disease resistance from *Brassica nigra* to *Brassica napus*. *Plant Breed.* 113, 106–113. doi: 10.1111/j.1439-0523.1994.tb00712.x

Glimelius, K. (1999). "Somatic hybridization," in *Biology of Brassica Coenospecies*, ed C. Gomez-Campo (Amsterdam: Elsevier), 107–148.

Goll, M. G., and Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514. doi: 10.1146/annurev.biochem.74.010904.153721

Ha, M., Lu, J., Tian, L., Ramachandran, V., Kasschau, K. D., Chapman, E. J., et al. (2009). Small RNAs serve as a genetic buffer against genomic shock in *Arabidopsis* interspecific hybrids and allopolyploids. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17835–17840. doi: 10.1073/pnas.0907003106

Ha, M., Ng, D. W., Li, W. H., and Chen, Z. J. (2011). Coordinated histone modifications are associated with gene expression variation within and between species. *Genome Res.* 21, 590–598. doi: 10.1101/gr.116467.110

Hansen, L. N., and Earle, E. D. (1994). Novel flowering and fatty acid characters in rapid cycling *Brassica napus* L. resynthesized by protoplast fusion. *Plant Cell Rep.* 14, 151–156. doi: 10.1007/bf00233780

Hansen, L. N., and Earle, E. D. (1995). Transfer of resistance to *Xanthomonas campestris* pv. *campestris* into *Brassica oleracea* L. by protoplast fusion. *Theor. Appl. Genet.* 91, 1293–1300. doi: 10.1007/BF00220944

Hansen, L. N., and Earle, E. D. (1997). Somatic hybrids between *Brassica oleracea* l. and *Sinapis alba* L. with resistance to Alternaria brassicae (Berk) Sacc. *Theor. Appl. Genet.* 94, 1078–1085. doi: 10.1007/s001220050518

Jauhar, P. P. (2006). Modern biotechnology as an integral supplement to conventional plant breeding: the prospects and challenges. *Crop Sci.* 46, 1841–1859. doi: 10.2135/cropsci2005.07-0223

Johnston, P. A., Timmerman-Vaughan, G. M., Farnden, K. J., and Pickering, R. (2009). Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. *Theor. Appl. Genet.* 118, 1429–1437. doi: 10.1007/s00122-009-0992-7

Jourdan, P., and Salazar, E. (1993). *Brassica carinata* resynthesized by protoplast fusion. *Theor. Appl. Genet.* 86, 567–572. doi: 10.1007/BF00838710

Kashkush, K., Feldman, M., and Levy, A. A. (2003). Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat. Genet.* 33, 102–106. doi: 10.1038/ng1063

Kraitshtein, Z., Yaakov, B., Khasdan, V., and Kashkush, K. (2010). Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics* 186, 801–812. doi: 10.1534/genetics.110.120790

Leflon, M., Brun, H., Eber, F., Delourme, R., Lucas, M. O., Vallée, P., et al. (2007). Detection, introgression and localization of genes conferring specific resistance to *Leptosphaeria maculans* from *Brassica rapa* into *B. napus*. *Theor. Appl. Genet.* 115, 897–906. doi: 10.1007/s00122-007-0616-z

Leflon, M., Eber, F., Letanneur, J. C., Chelysheva, L., Coriton, O., Huteau, V., et al. (2006). Pairing and recombination at meiosis of *Brassica rapa* (AA) and *Brassica napus* (AACC) hybrids. *Theor. Appl. Genet.* 113, 1467–1480. doi: 10.1007/s00122-006-0393-0

Liu, S., Li, F., Kong, L., Sun, Y., Qin, L., Chen, S., et al. (2015). Genetic and epigenetic changes in somatic hybrid introgression lines between wheat and tall wheatgrass. *Genetics* 199, 1035–1045. doi: 10.1534/genetics.114.174094

Liu, S. W., Zhao, S. Y., Chen, F. G., and Xia, G. M. (2007). Generation of novel high quality HMW-GS genes in two introgression lines of *Triticum aestivum*/*Agropyron elongatum*. *BMC Evol. Biol.* 7:76. doi: 10.1186/1471-2148-7-76

Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe *Brassiceae*. *Genome Res.* 15, 516–525. doi: 10.1101/gr.3531105

Madlung, A., Masuelli, R., Watson, B., Reynolds, S., Davison, J., and Comai, L. (2002). Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis allotetraploids*. *Plant Physiol.* 129, 733–746. doi: 10.1104/pp.003095

Menda, N., Strickler, S. R., Edwards, J. D., Bombarely, A., Dunham, D. M., Martin, G. B., et al. (2014). Analysis of wild-species introgressions in tomato inbreds uncovers ancestral origins. *BMC Plant Biol.* 28:287. doi: 10.1186/s12870-014-0287-2

Mennella, G., Rotino, G. L., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., et al. (2010). Characterization of health-related compounds in eggplant (*Solanum melongena* L.) lines derived from introgression of allied species. *J. Agric. Food Chem.* 58, 7597–7603. doi: 10.1021/jf101004z

Nagaharu, U. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Bot.* 7, 389–452.

Navrátilová, B., (2004). Protoplast cultures and protoplast fusion focused on Brassicaceae-a review. *Hortic. Sci.* 31, 140–157. Available online at: http://81.0.228.28/publicFiles/51556.pdf; http://www.agriculturejournals.cz/publicFiles/51556.pdf

Nicolas, S. D., Le Mignon, G., Eber, F., Coriton, O., Momod, H., Clouet, V., et al. (2007). Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids. *Genetics* 175, 487–503. doi: 10.1534/genetics.106.062968

Osborn, T. C., Butrulle, D. V., Sharpe, A. G., Pickering, K. J., Parkin, I. A., Parkin, I. A. P., et al. (2003). Detection and effects of a homeologous reciprocal transposition in *Brassica napus*. *Genetics* 165, 1569–1577.

Panjabi, P., Jagannath, A., Bisht, N. C., Padmaja, K. L., Sharma, S., Gupta, V., et al. (2008). Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using Intron Polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. *BMC Genomics* 9:113. doi: 10.1186/1471-2164-9-113

Paszkowski, J., and Whitham, S. A. (2001). Gene silencing and DNA methylation processes. *Curr. Opin. Plant Biol.* 4, 123–129. doi: 10.1016/S1369-5266(00)00147-3

Pires, J. C., Zhao, J. W., Schranz, M. E., Leon, E. J., Quijada, P. A., Lukens, L. N., et al. (2004). Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (*Brassicaceae*). *Biol. J. Linn. Soc.* 82, 675–688. doi: 10.1111/j.1095-8312.2004.00350.x

Primard-Brisset, C., Poupard, J. P., Horvais, R., Eber, F., Pelletier, G., Renard, M., et al. (2005). A new recombined double low restorer line for the Ogu-INRA cms in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 111, 736–746. doi: 10.1007/s00122-005-2059-8

Rangel, P. N., Brondani, R. P., Rangel, P. H., and Brondani, C. (2008). Agronomic and molecular characterization of introgression lines from the interspecific cross *Oryza sativa* (BG90-2) x *Oryza glumaepatula* (RS-16). *Genet. Mol. Res.* 7, 184–195. doi: 10.4238/vol7-1gmr406

Rieseberg, L. H., Sinervo, B., Linder, C. R., Ungerer, M. C., and Arias, D. M. (1996). Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. *Science* 272, 741–745. doi: 10.1126/science.272.5262.741

Roderick, H. W., Morgan, W. G., Harper, J. A., and Thomas, H. M. (2003). Introgression of crownrust (*Puccinia coronata*) resistance from meadow fescue (*Festuca pratensis*) into Italian rye grass (*Lolium multiflorum*) and physical mapping of the locus. *Heredity* 91, 396–400. doi: 10.1038/sj.hdy.6800344

Scholze, P., Kramer, R., Ryschka, U., Klocke, E., and Schumann, G. (2010). Somatic hybrids of vegetable brassicas as source for new resistances to fungal and virus diseases. *Euphytica* 176, 1–14. doi: 10.1007/s10681-010-0205-0

Shaked, H., Kashkush, K., Ozkan, H., Feldman, M., and Levy, A. A. (2001). Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* 13, 1749–1759. doi: 10.1105/tpc.13.8.1749

Tu, Y. Q., Sun, J., Liu, Y., Ge, X. H., Zhao, Z. G., Yao, X. C., et al. (2008). Production and characterization of intertribal somatic hybrids of *Raphanus sativus* and *Brassica rapa* with dye and medicinal plant *Isatis indigotica*. *Plant Cell Rep.* 27, 873–883. doi: 10.1007/s00299-008-0513-1

Udall, J. A., Quijada, P. A., and Osborn, T. C. (2005). Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of *Brassica napus L*. *Genetics* 169, 967–979. doi: 10.1534/genetics.104.033209

Wang, G., He, Q., Liu, F., Cheng, Z., Talbert, P. B., and Jin, W. (2011). Characterization of CENH3 proteins and centromere-associated DNA sequences in diploid and allotetraploid *Brassica* species. *Chromosoma* 120, 353–365. doi: 10.1007/s00412-011-0315-z

Wang, G. X., Tang, Y., Yan, H., Sheng, X. G., Hao, W. W., Zhang, L., et al. (2011). Production and characterization of interspecific somatic hybrids between *Brassica oleracea* var. *botrytis* and *B. nigra* and their progenies for the selection of advanced pre-breeding materials. *Plant Cell Rep.* 30, 1811–1821. doi: 10.1007/s00299-011-1088-9

Wang, Y. M., Dong, Z. Y., Zhang, Z. J., Lin, X. Y., Shen, Y., Zhou, D., et al. (2005). Extensive *de novo* genomic variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.). *Genetics* 170, 1945–1956. doi: 10.1534/genetics.105.040964

Wang, Y. P., Sonntag, K., and Rudloff, E. (2003). Development of rapeseed with high erucic acid content by asymmetric somatic hybridization between *Brassica napus* and *Crambe abyssinica*. *Theor. Appl. Genet.* 106, 1147–1155. doi: 10.1007/s00122-002-1176-x

Westman, A. L., Kresovich, S., and Dickson, M. H. (1999). Regional variation in *Brassica nigra* and other weedy crucifers for disease reaction to *Alternaria brassicicola* and *Xanthomonas campestris* pv. *campestris*. *Euphytica* 106, 253–259. doi: 10.1023/A:1003544025146

Wierzbicki, A. T., Haag, J. R., and Pikaard, C. S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell* 135, 635–648. doi: 10.1016/j.cell.2008.09.035

Xia, G. M. (2009). Progress of chromosome engineering mediated by asymmetric somatic hybridization. *J. Genet. Genomics* 36, 547–556. doi: 10.1016/S1673-8527(08)60146-0

Zhang, X., Ge, X., Shao, Y., Sun, G., and Li, Z. (2013). Genomic change, retrotransposon mobilization and extensive cytosine methylation alteration in *Brassica napus* introgressions from two intertribal hybridizations. *PLoS ONE* 8:e56346. doi: 10.1371/journal.pone.0056346

Zhao, N., Zhu, B., Li, M. J., Wang, L., Xu, L. Y., Zhang, H., et al. (2011). Extensive and heritable epigenetic remodeling and genetic stability accompany allohexaploidization of wheat. *Genetics* 188, 499–509. doi: 10.1534/genetics.111.127688

Zhong, X. B., Hans de Jong, J., and Zabel, P. (1996). Preparation of tomato meiotic pachytene and mitotic metaphase chromosomes suitable for fluorescence *in situ* hybridization (FISH). *Chromosome Res.* 4, 24–28. doi: 10.1007/BF02254940

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Wang, Lv, Zhang, Han, Zong, Guo, Zeng, Zhang, Wang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 November 2015
doi: 10.3389/fpls.2015.00957

Genome-specific differential gene expressions in resynthesized *Brassica* allotetraploids from pair-wise crosses of three cultivated diploids revealed by RNA-seq

Dawei Zhang[1], *Qi Pan*[1], *Cheng Cui*[2], *Chen Tan*[1], *Xianhong Ge*[1], *Yujiao Shao*[3]* and *Zaiyun Li*[1]*

[1] National Key Lab of Crop Genetic Improvement, National Center of Crop Molecular Breeding Technology, National Center of Oil Crop Improvement (Wuhan), College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [2] Crop Research Institute, Sichuan Academy of Agricultural Sciences, Chengdu, China, [3] College of Chemistry and Life Science, Hubei University of Education, Wuhan, China

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Genlou Sun,
Saint Mary's University, Canada
Maoteng Li,
Huazhong University of Science and Technology, China
Naghabushana K. Nayidu,
University of Saskatchewan, Canada

***Correspondence:**
Yujiao Shao
syjsyj520@126.com;
Zaiyun Li
lizaiyun@mail.hzau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: 31 July 2015
Accepted: 20 October 2015
Published: 04 November 2015

Citation:
Zhang D, Pan Q, Cui C, Tan C, Ge X, Shao Y and Li Z (2015) Genome-specific differential gene expressions in resynthesized Brassica allotetraploids from pair-wise crosses of three cultivated diploids revealed by RNA-seq. Front. Plant Sci. 6:957. doi: 10.3389/fpls.2015.00957

Polyploidy is popular for the speciation of angiosperms but the initial stage of allopolyploidization resulting from interspecific hybridization and genome duplication is associated with different extents of changes in genome structure and gene expressions. Herein, the transcriptomes detected by RNA-seq in resynthesized *Brassica* allotetraploids (*Brassica juncea*, AABB; *B. napus*, AACC; *B. carinata*, BBCC) from the pair-wise crosses of the same three diploids (*B. rapa*, AA; *B. nigra*, BB; *B. oleracea*, CC) were compared to reveal the patterns of gene expressions from progenitor genomes and the effects of different types of genome combinations and cytoplasm, upon the genome merger and duplication. From transcriptomic analyses for leaves and silique walls, extensive expression alterations were revealed in these resynthesized allotetraploids relative to their diploid progenitors, as well as during the transition from vegetative to reproductive development, for differential and transgressive gene expressions were variable in numbers and functions. Genes involved in glucosinolates and DNA methylation were transgressively up-regulated among most samples, suggesting that gene expression regulation was immediately established after allopolyploidization. The expression of ribosomal protein genes was also tissue-specific and showed a similar expression hierarchy of rRNA genes. The balance between the co-up and co-down regulation was observed between reciprocal *B. napus* with different types of the cytoplasm. Our results suggested that gene expression changes occurred after initial genome merger and such profound alterations might enhance the growth vigor and adaptability of *Brassica* allotetraploids.

Keywords: *Brassica* species, allopolyploidization, differential gene expressions, transgressive gene expression, transcriptome

INTRODUCTION

Allopolyploidization which is realized through the merger and duplication of distinct parental genomes following interspecific hybridizations of two or more related species results in the origin of new allopolyploid species (Otto, 2007; Doyle et al., 2008). This pattern of speciation occurs widely in angiosperms, largely because the allopolyploids relative to their progenitors show the enhanced growth vigor (the phenomenon of heterosis) and the advantage in ecological adaptation (Chen, 2007; Leitch and Leitch, 2008). The obvious success of allopolyploidy in nature has allured the extensive investigations of genetic effects caused by genome merger at levels of the chromosomes, DNA sequences, gene expression, proteins, small RNA, and DNA methylation during the last two decades, by utilizing the continuously improved approaches of molecular biology and genome sequencing (Song et al., 1995; Soltis and Soltis, 2012; Li et al., 2014; Soltis et al., 2014). The results from recent and synthetic allopolyploids of different taxa (*Arabidopsis*, *Brassica*, cotton, wheat, triticale, etc.) demonstrate that the initial stage of allopolyploidization is accompanied by the various genetic, epigenetic and transcriptome changes, but the degrees of variations are obviously different between allopolyploids (Soltis and Soltis, 2012). So the new allopolyploids respond to the genome merger by rapid alterations in genomic structure and function, in order to coordinate the divergent genome at different aspects and to establish the novel plants for further evolution. As to the crucial gene expression in hybrids and new allopolyploids, widespread changes are revealed by transcriptomic analysis and extensive gene expression changes are non-additive in allopolyploids relative to their parents, including expression level dominance and transgression (outside the range of either parent), such profound changes to gene expression may enable new hybrids to survive in novel environments not accessible to their parent species (Wang et al., 2006; Chen, 2007; Rapp et al., 2009; Flagel and Wendel, 2010; Yoo et al., 2013; Li et al., 2014).

The six cultivated *Brassica* species offer a text book system of allopolyploidization through the pair-wise crosses of three diploids, which is illustrated as U-triangle (U 1935). *Brassica carinata* ($2n = 34$, BBCC), *B. juncea* ($2n = 36$, AABB), and *B. napus* ($2n = 38$, AACC) are three allotetraploids, which are derived from various two-way combinations of the three diploids *B. nigra* ($2n = 16$, BB), *B. oleracea* ($2n = 18$, CC), and *B. rapa* ($2n = 20$, AA). Resynthesized *Brassica* allotetraploids through the interspecific hybridizations between three extant *Brassica* diploid progenitors have been widely investigated to elucidate the genetic alterations during the initial stage of allopolyploid formation, since the seminal work of Song et al. (1995). The reciprocal synthetics of allotetraploids would remedy the drawback of unprecise progenitors for natural counterparts and also uniparental cytoplasm background. The results obtained provide many new insights into the genome regulations following genome merger, though the studies focused mainly on the younger *B. napus* with much more agricultural and economic importance (Gaeta et al., 2007; Nicolas et al., 2007, 2012; Xu et al., 2009; Szadkowski et al., 2010, 2011; Xiong et al., 2011; Cui et al., 2012, 2013; Ge et al., 2013). The genome of *B. rapa* has been sequenced, which should enhance the evolutionary analysis of these *Brassica* allotetraploids (Wang et al., 2011). By using the *B. rapa* as reference genome, the analysis on resynthesized *B. napus* across four generations revealed that the gene expression was more complicated than the simple combination of two genomes, and non-additive gene regulation was also detected (Jiang et al., 2013). Another transcriptomic study on synthesized *Brassica* allohexaploid and its parents showed that genome-wide changes in gene expression were involved in adaptation and evolution processes, and non-additive genes associated with important biological processes were identified (Zhao et al., 2013). An intriguing genetic interaction in three *Brassica* allopolyploids was the hierarchy of nucleolar dominance (genomes BB> AA> CC) in three allotetraploids, in which both *B. juncea* (BB> AA) and *B. carinata* (BB> CC) expressed the rRNA genes from *B. nigra*, and *B. napus* (AA> CC) expressed those from *B. rapa*, but the genes from another parent were silenced (Chen and Pikaard, 1997; Ge et al., 2013). Meanwhile, the rRNA genes silenced in vegetative tissues were expressed in reproductive tissues, indicated that the expression of rRNA genes could be tissue-specific (Chen and Pikaard, 1997). However, whether the parent-specific expression of the genes for ribosomal protein also occurs is still an open question.

In this study, the gene expressions detected by RNA-seq in three resynthesized *Brassica* allotetraploids prior to meiosis from the pair-wise crosses of the same three diploids are compared, with the aims: (1) to reveal the differential gene expressions from progenitor genomes after genome merger and duplication and with the exclusion of meiotic homoeologous exchanges; (2) to assess transgressive gene expression and its effect on adaption environment changes in *Brassica* allotetraploids; (3) to identify the expression of ribosomal protein genes after genome merger and the relationship with the nucleolar dominance. The results should give new insights into the genetic contribution of three progenitors and their interactions at the beginning of *Brassica* allopolyploidization.

MATERIALS AND METHODS

Plant Materials and Sample Preparations

Four *Brassica* allotetraploids (AACC/CCAA, AABB, and BBCC) used in this study were previously synthesized from pair-wise crosses of three cultivated diploids, *B. rapa* L. ($2n = 20$, AA genome, genotype 3H120), *B. nigra* (L.) Koch cv. Giebra ($2n = 16$, BB), *B. oleracea* var. *alboglabra* L. ($2n = 18$, CC, genotype Chi Jie Lan), with the aid of embryo rescue and colchicine inducing chromosome doubling *in vitro* (Cui et al., 2012, 2013). Three allotetraploids (AABB, AACC, and BBCC) were synthesized by doubling the chromosome numbers of the respective digenomic hybrids, and CCAA was directly obtained from the cultured embryo derived plantlets, probably by the spontaneous chromosome doubling *in vitro*. The plants of each synthetic were derived from single hybrid embryo by successive

Abbreviations: DEGs, differentially expressed genes; FC, fold change; GO, gene ontology.

subculturing on MS medium, and those of each parent were derived from one seed. The rooted plantlets of four allotetraploids and three parents cultured on the medium were transferred to the pots and kept in the unheated greenhouse. Leaves at the same stage were collected from these seven materials, while the silique walls after 21 days of pollination were collected. All samples were stored in liquid nitrogen and kept at −80 until RNA extraction.

RNA-seq Library Preparation and Sequencing

Total RNAs were extracted by TRIzol reagent (invitrogen) and were then treated with RQ1 DNase (promega) to remove DNA. The quality and quantity of the purified RNA were determined by measuring the absorbance at 260/280 nm (A260/A280) using smartspec plus (BioRad). RNA integrity was further verified by 1.5% agrose gel electrophoresis.

For each sample, 10 μg of total RNA was used for RNA-seq library preparation. Polyadenylated mRNAs were purified and concentrated with oligo(dT)-conjugated magnetic beads (invitrogen) before used for directional RNA-seq library preparation. Purified mRNAs were iron fragmented at 95°C followed by end repair and 5′ adaptor ligation. Then reverse transcription was performed with RT primer harboring 3′ adaptor sequence and randomized hexamer. The cDNAs were purified and amplified and PCR products corresponding to 200–500 bps were purified, quantified, and stored at −80°C until used for sequencing.

For high-throughput sequencing, the libraries were prepared following the manufacturer's instructions and applied to illumina GAIIx system for 80 nt single-end sequencing by ABlife. Inc (Wuhan, China). Owing to the high cost for sequencing 2 years ago, the total of 14 (2 tissues × 7 plant samples) libraries without replicates were made for RNA-seq.

Reads Filtering and Alignment

After the high-throughput sequencing, the raw data which contained adapter and low quality cycle of reads were filtered. Due to lack of reference genome from *B. nigra* and the genome sequence of *B. oleracea* was still unpublished yet 2 years ago (Liu et al., 2014), the clean reads were aligned to the *B.rapa var. pekinensis Chiifu-401* genome as described in previse study (Jiang et al., 2013; Zhao et al., 2013) using the default parameters for TopHat, allowing no more than two mismatched bases. Subsequently, only unique mapped reads were used in further study to provide sensitive and accurate alignment results, even for highly repetitive genomes. The gene expression level was calculated by using RPKM method (Reads Per kb per Million reads). If there were more than one transcript for a gene, the longest one was used to calculate its expression level and coverage.

Analysis of Differentially Expressed Genes (DEGs)

Program edgerR was used to analyze the differentially expressed genes between samples. Fold change ≥ 2 and $P \leq 0.01$ between compared samples were considered as simulated biological variation in DEG analysis. GO enrichment was performed using blast2Go (http://www.blast2go.com/b2ghome). In addition, the web-based *Brassica* database (http://brassicadb.org/brad/index.php) and WEGO (Ye et al., 2006) server for gene ontology analysis was also used in this study ($P < 0.05$). We also used DAVID to investigate the transgressively expressed genes, GO terms with enrichment score ≥ 0.5 and $P < 0.05$ were considered significantly enriched (Huang et al., 2008).

Pearson correlation between biological samples was calculated using IBM SPSS V19.0 software. Additionally, the Venn diagrams in this study were performed using online tool (http://bioinfogp.cnb.csic.es/tools/venny/).

Analysis of r-protein Genes

All the mapped genes were blasted to the *Arabidopsis* genome, those genes matching the *Arabidopsis* ribosomal protein genes were considered to be r-protein genes in *Brassica*. To group r-protein genes with similar expression patterns, a hierarchical clustering was generated using the normalized expression values ($\log_2$ RPKM) from each library. The analysis was conducted using HemI software with Pearson correlation as the distance measure (Deng et al., 2014).

Real-time Quantitative RT-PCR (qRT-PCR) Analysis

The RNA samples used for the qRT-PCR assays were the same as for the RNA-seq experiments. First-strand cDNA synthesis was performed with 1500 ng of total RNA using Thermo Scientific RevertAid First Strand cDNA Synthesis Kit, total RNA (0.5 μg) was reverse-transcribed with oligo (dT)18 primer (0.5 μg/μl) according to the described protocol. Gene-specific primers were designed according to the reference unigene sequences using the Primer 3.0, all primer sequences are listed in **Supplementary Table 5**. A primer was also designed for *B. napus* actin gene to normalize the amplification efficiency. QRT-PCR assays in triplicate were performed using Kapa Probe Fast qPCR Kit with a Bio-Rad CFX96 Real-Time Detection System. The actin gene was used as an internal control for data normalization, and quantitative variation in the different replicates was calculated using the delta-delta threshold cycle relative quantification method.

Availability of Supporting Data

The datasets supporting the results of this article are available in GenBank SRA under accession ID PRJNA281555. Other supporting data are included within the article and its additional files.

RESULTS

Gene Expressions between Nascent Allotetraploids and Parents

To investigate and compare the mRNA expression levels in the resynthesized *Brassica* allotetraploids relative to their diploid progenitors (**Figure 1**), 14 RNA-seq libraries were constructed for two types of tissue: leaves (L. for short) and silique walls (S.),

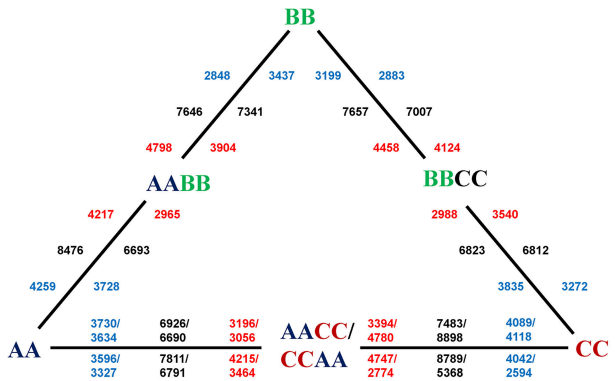


FIGURE 1 | DEGs between synthetic allotetraploids and their diploid parents. The numbers in the inner circle of the triangle are DEGs in leaves, and those on the periphery are DEGs in silique walls. The total number (black) of DEGs between each allotetraploid and one of parent is in the middle, the number of up-regulation (red) is close to the allotetraploid, and the number of down-regulation (blue) is near to the parent. The lower number of DEGs for the silique sample of CCAA might be caused by the fewer raw reads for the sequencing problems.

respectively. As a result, we obtained an average of 12,748,770 (84.02%) high-quality and clean reads from the raw reads (**Supplementary Table 1**). Among the clean reads, the average 41.76% reads were matched either to unique or multiple genomic positions using the *B. rapa* reference genome. As most of uniquely mapped reads were mature mRNA or ncRNA and multiple mapped reads were mainly rRNA and tRNA, only uniquely and perfectly mapped reads (3,696,642) were used to measure the transcriptional activity of each gene. Finally, our RNA-Seq data revealed an average of 28944 genes in *Brassica* allotetraploids and its progenitors, accounting for 70.56% of the total genes in *B. rapa* reference genome. For further comparative analysis, the gene expression level was calculated using RPKM method (Reads Per kb per Million reads).

A correlation dendrogram showed that the gene expression patterns were distinct between leaves and silique walls, for the samples of the two tissues were clustered separately (**Figure 2**). However, the global relationships of gene expression among the synthetics and their parents were the same in two tissues. In the dendrogram, AACC and CCAA which were most closely correlated were clustered firstly with AA and then with AABB, and BBCC was clustered with CC, while BB appeared as the outlier.

Differentially Expressed Genes (DEGs) in Nascent Allotetraploids

To study the gene expression patterns during allopolyploidization process, we first performed pair-wise comparisons between allotetraploids to their parents to identify differentially expressed genes (DEGs) using edgerR (fold change ≥ 2 and $P \leq 0.01$ as criteria; **Supplementary Figure 1**). As a result, average 3791 (13.1% of expressed genes) DEGs were up-regulated and 3534(12.2% of expressed genes) DEGs were down-regulated in allotetraploids, respectively (**Figure 1**). There was no significant difference between the average number of up and down-regulated genes in leaves (3593 vs. 3721; *t*-test, $P > 0.05$), however, the difference was statistically significant (3990 vs. 3347; *t*-test, $P < 0.05$) in silique walls, suggesting that the direction of DEGs was affected by tissue type. The maximum number (8898) of DEGs was observed between L.CCAA and L.CC among all comparisons, including 4780 up-regulated and 4118 down-regulated. But the minimum number (5368) of DEGs was between S.CCAA and S.CC, only 2774 genes were up-regulated and 2594 were down-regulated.

Notably, by comparing the total number of genes showing differential expression (both up and down regulation) between the allotetraploids and their parents, there was a bias in the direction of differential expression relative to the parents (**Table 1**). For example, more expressed genes remained statistically unchanged (less differential expression) between L.AACC/CCAA and L.AA than between L.AACC/CCAA and L.CC (Chi square test, $P < 0.01$). This asymmetric gene expression was also observed in L.AABB and L.BBCC, the global expression patterns were closer to either L.AA in L.AABB or L.CC in L.BBCC (Chi square test, $P < 0.01$). Whereas, the expression bias in silique walls was not as tangible as in leaves, the global expression patterns were closer to either S.BB in S.AABB or S.CC in L.AACC.

Identification of Transgressively Expressed Genes

Among the differentially expressed genes, we then filtered the transgressively expressed genes in allotetraploids which showed more than two-fold changes in expression relative to both parents (**Table 2**). Briefly, on average 878 genes (3.0% of total expressed genes) were transgressively up-regulated while 652 genes (2.3% of total expressed genes) were transgressively down-regulated. When we compared the number of genes for up and down-regulation in each sample, no significant bias was observed in leaves, but there were more genes exhibiting transgressive upregulation in silique walls.

Moreover, the comparisons of transgressively expressed genes in each *Brassica* allotetraploids revealed that transgressive expression varied between tissues (**Figure 3**). For instance, more genes showed transgressive up-regulation in silique walls than leaves in AABB (1172 vs. 566). Among those genes, a multitude of them were specifically expressed in different direction and tissues, suggesting that the majority of transgressively expressed genes were tissue-specific. In addition, certain number of genes showed co-transgressive up- or down-regulation (105 and 40) regardless of tissue type, likely due to genome merger. However, 28 genes showing transgressive down-regulation in leaves displayed up-regulation in silique walls, while only nine genes showed expression changes in the opposite direction (From up in leaves to down in silique walls; **Figure 3**). The similar tendency of gene expression changes was observed in other *Brassica* allotetraploids (28 vs. 6 in BBCC; 21 vs. 5 in CCAA; 28 vs. 9 in AACC), suggesting that genes in silique walls tended to be transgressively up-regulated.

Transgressively expressed genes were further functionally classified according to Gene Ontology (GO) terms using

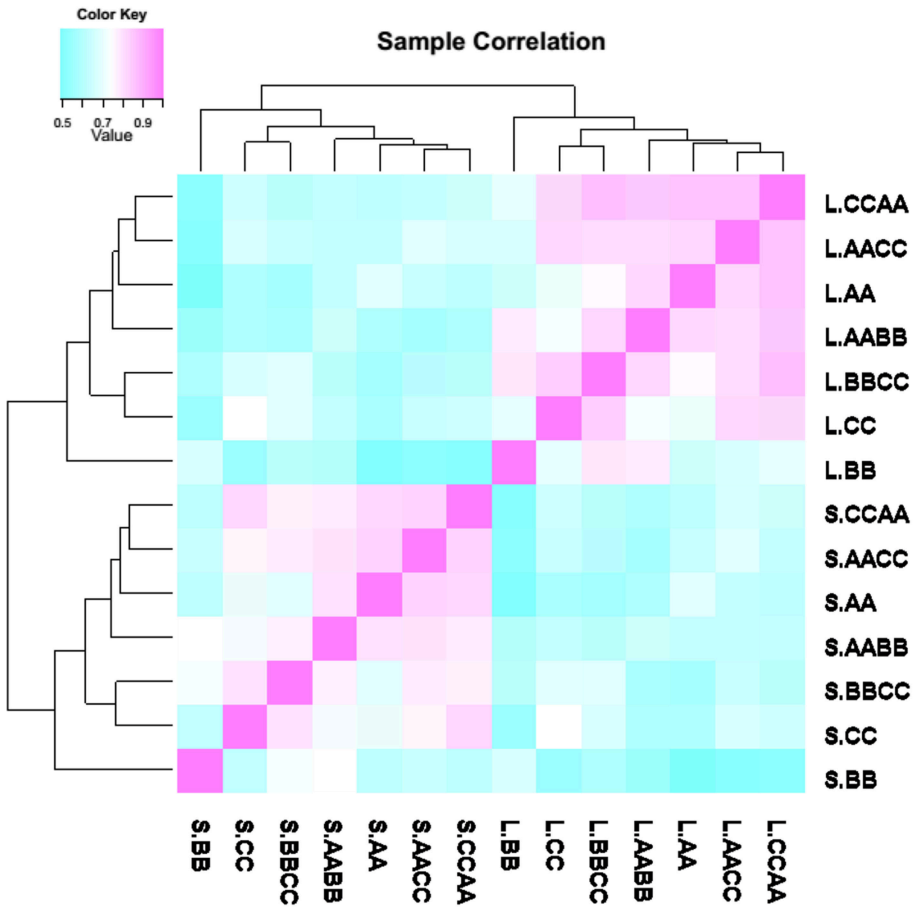


FIGURE 2 | Gene expression correlations between nascent allotetraploids and parents. The branch length represents the extent of gene expression variance among samples.

DAVID (Huang et al., 2008). Ten functional clusters with the highest enrichment score were selected with criteria that enrichment score > 0.5 and $P < 0.05$(see more details in **Supplementary Table 2**). Although remarkable difference was observed among samples, it was interesting to note that a certain functional cluster encoding glucosinolates was up-regulated among a group of allotetraploids (L.BBCC, L.CCAA, L.AACC, S.AABB, S.AACC), suggesting that the increased expression of resistance-related genes, especially glucosinolates, was common in *Brassica* allotetraploids (**Supplementary Table 2**).

Co-transgressive Gene Expression in *Brassica* Allotetraploids

To further ascertain which genes were specially expressed in tissues regardless of genome composition and what the function of those genes was, we investigated genes showing co-transgressive expression in leaves and silique walls, respectively. To avoid confusion, three representative allotetraploids (AABB, BBCC, and CCAA) with different cytoplasm were selected for further comparative analysis.

As illustrated in the Venn diagrams, 17 genes exhibited co-transgressive up-regulation and 18 genes gave co-transgressive down-regulation among the allotetraploids in leaves, respectively (**Figures 4A,B**). Interestingly, two key methylase genes (Bra002610 and Bra022537) which were involved in DNA methylation on cytosine and methyltransferase activity reflected

TABLE 1 | Summary of differential expression between the *Brassica* allotetraploids and their parents.

	Maternal parent[a]	Paternal parent	Less differential expression genome
L.AABB	6693	7341	A*
L.BBCC	7657	6823	C*
L.CCAA	8898	6687	A*
L.AACC	6926	7483	A*
S.AABB	8476	7646	B*
S.BBCC	7007	6812	C
S.CCAA	5368	6791	C*
S.AACC	7098	8019	A*

[a] *The total number of genes showing differential expression between the allotetraploids and their maternal parent, including both up and down-regulated genes.*
**Chi square test with expected ration of 1:1, $P < 0.01$.*

co-transgressive up-regulation in the allotetraploids, suggesting that methylation-related genes were up-regulated during *Brassica* allopolyploidization in leaves (**Supplementary Table 4**). In silique walls, there were more genes displaying co-transgressive up-regulation than down-regulation (23 vs. 5, **Figures 4C,D**).

TABLE 2 | Transgressive expressions in *Brassica* allotetraploids.

	Up-regulation (%)[a]	Down-regulation (%)	Biased regulation
L.AABB	566(2.0)	489(1.8)	Up
L.BBCC	828(2.9)	841(2.9)	Down
L.CCAA	676(2.3)	692(2.3)	Down
L.AACC	649(2.3)	957(3.3)	Down*
S.AABB	1172(4.0)	467(1.6)	Up*
S.BBCC	1021(3.4)	391(1.3)	Up*
S.CCAA	647(2.3)	482(1.7)	Up*
S.AACC	1469(4.8)	896(2.9)	Up*
Average	878(3.0)	652(2.3)	Up*

[a]*Calculated by dividing the number of expressed genes in each Brassica allotetraploid.*
**Chi square test with expected ration of 1:1, P < 0.01.*

Novel Gene Expression and Silencing in Allotetraploids

To explore novel gene expression and gene silencing in the allotetraploids, a strict criterion was set: novel expression was defined when both parental lines had no reads but in allotetraploids and the genes displayed more than 10 reads and RPKM $\geq$ 2. On the contrary, if both parental lines contained genes with more than 10 reads and RPKM $\geq$ 2, yet allotetraploids had no reads, we considered this situation as gene silencing.

Overall, a total of 160 genes were found to be novel expression, while 102 genes exhibited silencing in allotetraploids, suggesting that these changes were tightly controlled (**Table 3**). There was no significant difference between the total number of genes showing novel expression and silencing in both tissues, but considerable variation between the two different expression patterns was found in silique walls: more genes exhibited silencing than novel expression in S.CCAA, whereas more genes showed novel expression than silencing in S.BBCC and S.AABB. Furthermore, the number of genes showing novel expression in each sample was negatively correlated with those genes being silenced (Pearson correlation, $r = -0.92$, $P = 0.008 < 0.01$; **Figure 5**).

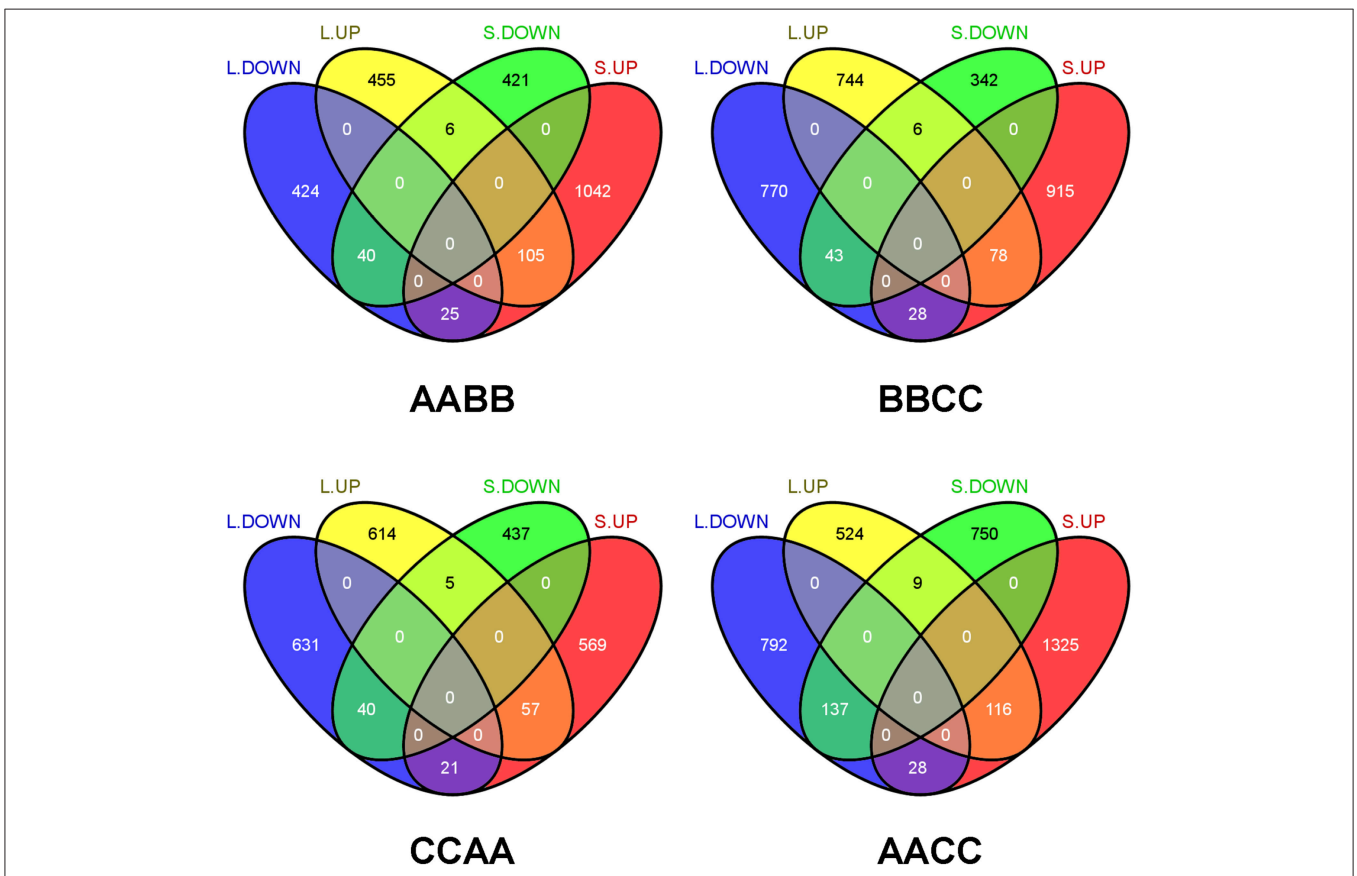


FIGURE 3 | Comparison of transgressive expression between leaves and silique walls in each *Brassica* allotetraploid. L.UP and L.DOWN: up- and down-regulation in leaves. S.UP and S.DOWN: up-/down regulation in silique walls.

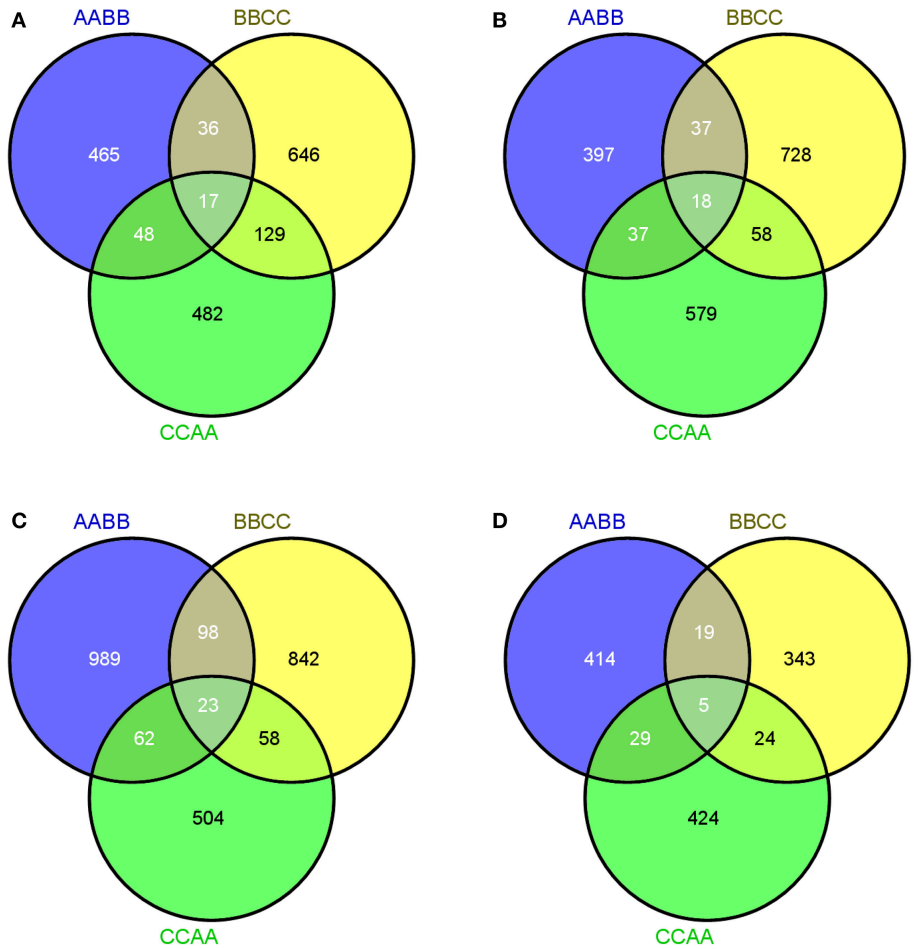


FIGURE 4 | Co-transgressive gene expressions of *Brassica* allotetraploids. (A,B) Up- and down-regulated genes in leaves. **(C,D)** Up- and down-regulated genes in silique walls.

The biological functions and processes of the two expression patterns were analyzed using gene ontology annotations according to function annotation convention (**Supplementary Figure 2**). The enriched GO terms were similar between tissues, while the number of genes was various in the enriched GO terms. The genes associated with cell part, cellular, as well as metabolic process and response to stimulus were over-represented in both tissues. However, there were more genes showing novel expression than silencing in those GO terms in leaves, but more genes exhibited silencing than novel expression in silique walls.

Differentially Expressed Genes between Reciprocal *B. napus*

To evaluate how gene expression was affected by cytoplasm types, we examined gene expression changes between the reciprocal synthetics of *B. napus* (AACC/CCAA) which kept the same nuclear genomes but the cytoplasm from *B. rapa* or *B. oleracea*. Though a large proportion of genes showed differential expression, the total number of up-regulated genes was similar to that of down-regulated genes (6210 vs. 6270), suggesting that there was no bias in the direction of differential expression between AACC and CCAA (Chi square text,$\chi^2 = 0.29$, $P > 0.01$; **Table 4**). Among these differentially expressed genes, only those showing co-regulation were selected for subsequent function analysis to avoid the tissues specific effects, the difference between co-up (581) and co-down (530) regulation was also insignificant (Chi square text, $\chi^2 = 2.25$, $P > 0.01$; **Table 4**).

To have an overview of the major differences between the reciprocal synthetics, we performed GO analysis of the differential expressed genes. We found that the number of co-up regulated genes was similar to that of co-down regulated in the most enriched classes (**Figure 6**). Among the cellular component categories, cell and cell parts were the largest enrichment, followed by organelle. As to the molecular function categories, the most over-represented GO terms were binding and catalytic, far more than others. In the biological process categories, there were more GO terms compared to the other two categories, and slightly more genes involved in the biological regulation, cellular process, metabolic process as well as response to stimulus were found to be up-regulated in these GO terms.

TABLE 3 | Genes showing novel expression and silencing in three *Brassica* allotetraploids.

Samples	Novel expression(%)[a]	Silencing(%)[b]
L.AABB	18(0.06)	27(0.10)
L.BBCC	18(0.06)	16(0.06)
L.CCAA	33(0.12)	10(0.04)
S.AABB	39(0.13)	5(0.01)
S.BBCC	42(0.14)	4(0.01)
S.CCAA	10(0.04)	40(0.14)
Total	160	102

[a,b]*Calculated by dividing the number of expressed genes in each Brassica allotetraploid.*

Analysis of r-protein Gene Expression

Besides the afore-mentioned expression patterns, we further identified ribosomal protein genes among samples to evaluate whether the genome merger has similar effect on housekeeping genes. We compared the *Arabidopsis* r-proteins against the above mapped genes based on the sequencing data. 363 genes were identified in our libraries which matched a total of 79 of the 80 r-proteins identified in *Arabidopsis,* suggesting that there was a high number of homologous r-proteins genes between *Brassica* and *Arabidopsis* (**Supplementary Table 4**; Whittle and Krochko, 2009). However, the expression of r-protein genes varied extensively across 14 libraries, the average expression level of r-protein genes in leaves was generally higher than that in silique walls (**Supplementary Figure 3**). The highest expression level of r-proteins genes were detected in L.AABB and L.BBCC, however, it was notable that the lowest expression level were also found in their silique walls (S.AABB and S.BBCC).

To further reveal the relationships of r-proteins gene expression among *Brassica* allotetraploids and diploids in different tissue, hierarchical clustering of expression level (standardized by $\log_2$ RPKM) was conducted for the 363 expressed r-proteins genes (encoding 79 r-proteins) using HemI (Deng et al., 2014). Those *Brassica* allotetraploids and diploids with similar expression profiles of r-proteins genes were more closely clustered in leaves and silique walls, respectively (**Figures 7**, **8**). In leaves, two allotetraploids with B-genome (L.AABB and L.BBCC) were closer to each other and clustered with L.BB, while reciprocal *B. napus* (L.AACC/L.CCAA) exhibited similar transcription profiles with L.AA and diverged markedly from L.CC. In silique walls, r-protein genes retained similar expression pattern to that in leaves on whole genome scale. Three allotetraploids with A-genome (S.AACC, S.CCAA, and S.AABB) were closer to S.AA and separated from S.CC and S.BBCC, while all of them were quite distinct from S.BB.

Verification of DEGs by qRT-PCR

To confirm that the above results of differentially expressed genes, a set of gene-specific primers were designed for quantitative RT-PCR assays (**Supplementary Table 5**). The relative transcript levels in the allotetraploids and their parents were compared with those of RNA-seq data (RPKM value). For 24 out of the 30 comparisons, qRT-PCR analysis revealed the same expression trends as the RNA-seq data, despite some quantitative differences (**Supplementary Figure 4**). Among 25 genes showing transgressive expression in the allotetraploids relative to their parents using RNA-Seq, 12 were up-regulated, eight were down-regulated, and five were differentially expressed to only one of their parents, respectively. Moreover, four of five genes showing differential expression between the reciprocal synthetics of *B. napus* also showed the same expression trends revealed by qRT-PCR and RNA-seq, confirming the reliability of RNA-seq data.

TABLE 4 | Genes showing co-regulation in leaves and silique walls between AACC and CCAA.

	Up-regulation	Down-regulation	Biased regulation
Leaves	2633	3285	Down*
Silique walls	3577	2985	Up*
Total	6210	6270	Up
Co-regulation	581(9.36%)	530(8.45%)	Up

**Chi square test with expected ration of 1:1, P < 0.01.*

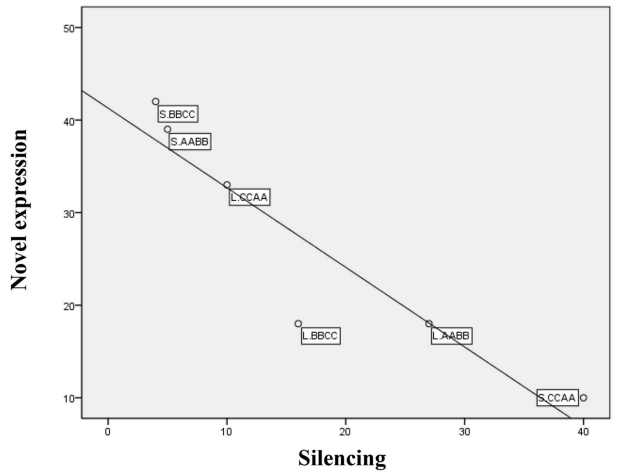


FIGURE 5 | Inverse correlations between number of genes for novel expression and silencing among samples.

DISCUSSION

Widespread changes of gene expression during allopolyploidization have been revealed in different aspects, including homoeolog expression bias, novel gene expression, or silencing, transgressive up or down-regulation, expression level dominance, and altered expression times and locations (Doyle et al., 2008). In this study, global transcriptome analyses through RNA-Seq were made for serial synthetic *Brassica* allotetraploids derived from the same three diploid parents (**Figure 1**; Cui et al., 2012). We focused on general analysis of transgressive gene expression, novel gene expression, and gene silencing, r-protein genes expression upon *Brassica* allopolyploidization. In addition, qRT-PCR analysis was also performed to confirm the reliability of RNA-seq data.

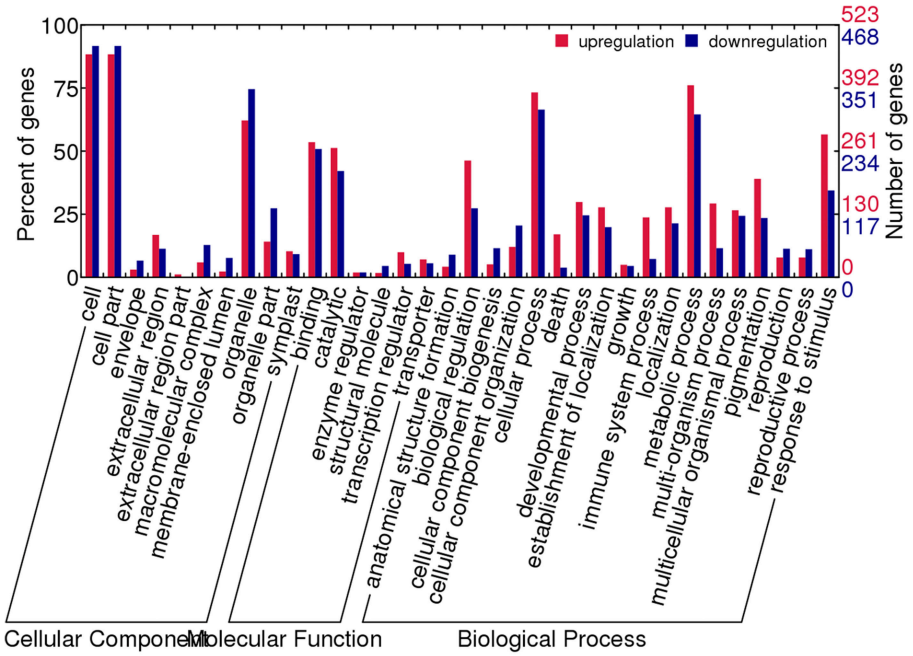


FIGURE 6 | GO functional categories of genes showing co-differential expression between AACC and CCAA.

Effects of Transgressive Gene Expression during Allopolyploidization

In consistence with other findings of gene expressions in cotton (Yoo et al., 2013), wheat (Li et al., 2014), and *Senecio* (Hegarty et al., 2008), transgressive gene expression in our *Brassica* allotetraploids was genome-wide and temporal. A lot of genes exhibited opposite expression direction between tissues, suggesting that transgressively expressed genes had different roles in development (**Table 2**). Additionally, transgressively up-regulated gene expression increased over time, more genes were or reversed to be transgressively up-regulated in silique walls (**Figure 3**). Moreover, wide ranges of alterations occurred for the functional clusters of transgressively expressed genes between tissues (**Supplementary Table 2**). All these changes indicated that transgressive gene expression was sophisticated in different genome background and different growth phases. In particular, transgressive up-regulation of resistance-related genes including for the synthesis of glucosinolates could be responsible for immediate physiological pre-adaptation of *Brassica* allotetraploids, since glucosinolates and their breakdown products played a role in the defense of plants against pathogen (Brader et al., 2001), fungi (Hiruma et al., 2010), and insects (Ahuja et al., 2015).

Interestingly, the genes which involved in DNA methylation were transgressively up-regulated in leaves across three types of *Brassica* allotetraploid (**Figure 4A**; **Supplementary Table 3**). DNA methylation was an important and best-studied epigenetic phenomenon in allopolyploids. Allopolyploids (e.g., wheat, *Senecio*, *Spartina*) underwent rapid and widespread metylation state changes and significant proportion of these changes showed non-additivity in previous studies (Dong et al., 2005; Salmon et al., 2005; Hegarty et al., 2011). Recent investigations in synthesized *Brassica* hybrids and allotetraploids also provided evidence for rapid changes in DNA methylation by transcriptome analysis or genetic analysis (Xu et al., 2012; Cui et al., 2013; Jiang et al., 2013; Zhao et al., 2013). Consistent with these studies, our data further demonstrated that the expression of methylation-related genes was non-additive and indeed transgressively up-regulated. DNA methylation was vitally important in silencing transposons and regulating gene expression, typically reducing expression (Zilberman et al., 2007; Zhang, 2008), as a consequence, the reduced expression and transposon silencing could prove capacity for adaptation.

Together, these data suggested that transgressive gene expression could be various in number and function between tissues. Such variation (including DNA methylation) might play a causative role in regulation of gene expressions in allopolyploids and increase their relative fitness over their parents in novel environment.

Negative Correlation between Novel Gene Expression and Silencing

Evidence for rapid gene silencing and novel expression came from the detection of missing parental fragments by cDNA - AFLP screens and verification by RT-PCR in *Arabidopsis* and cotton, and these expression changes could be variable in different parts of the plants (Adams et al., 2003; Wang et al., 2004). Recently, reciprocal gene silencing and novel expression were also found both in natural and synthetic cottons by using RNA-Seq, and specifically increased in natural allopolyploids (Yoo et al., 2013).



FIGURE 7 | Hierarchical clustering of r-protein genes among allotetraploids and diploids in leaves.

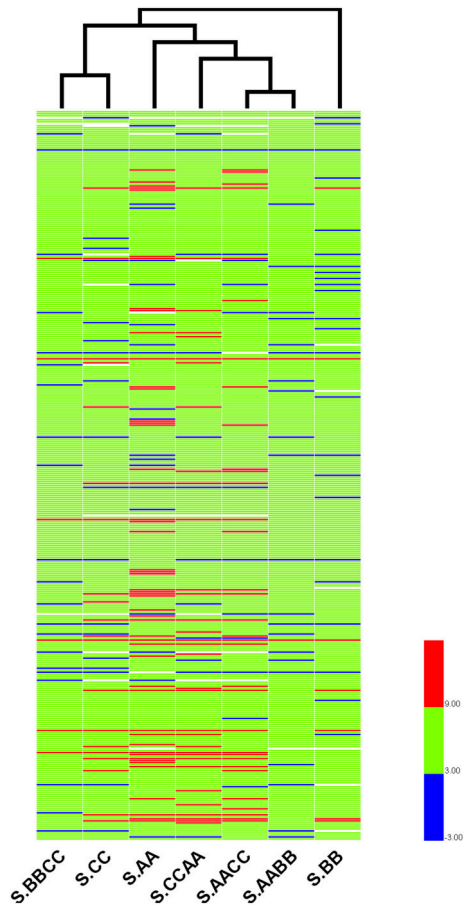


FIGURE 8 | Hierarchical clustering of r-protein genes among allotetraploids and diploids in silique walls.

To explore patterns of novel gene expression and silencing during the *Brassica* allopolyploidization, synthetic *Brassica* allotetraploids (U 1935) were examined using RNA-Seq data, since RNA-Seq technology could provide more data than cDNA -AFLP and RT-PCR. By restrictive criteria, the percentage of genes showing novel expression and silencing was significantly less than that of *Arabidopsis* and cotton using AFLP-cDNA and RT-PCR method (Adams et al., 2003; Wang et al., 2004), while similar to that of cotton from RNA-Seq (Yoo et al., 2013). This difference may be explained by technical consideration and various restrictive criteria. As a check in present study, we have also validated of five genes using qRT-PCR, four genes and one gene, respectively, were confirmed as novel expression and silencing (**Supplementary Figure 4**). Our results also found a wide range of variations between the two different expression patterns and different tissues of the allotetraploids (**Table 4**). More genes displayed novel expression or silencing in silique walls than in leaves, suggesting that tissue-specific expression partitioning could arise quickly after the onset of allotetraploids formation, which was consistent with the previous study (Adams et al., 2003; Wang et al., 2004; Adams and Wendel, 2005; Buggs et al., 2009). Interestingly, there was no significant difference between the total number of genes, but inverse correlation between the two expression patterns was observed among samples, indicating that there was a tradeoff between novel gene expression and silencing during plant development (**Figure 5**).

Cytoplasmic Effects on Genes Expression in Reciprocal *B. napus*

The cytoplasm has exhibited considerable influence on the evolution of nuclear genomes of allopolyploids (Prakash et al., 2009), because the presence of the paternal nuclear genome in the maternal cytoplasm could result in nuclear-cytoplasmic incompatibilities. Extensive and rapid genomic changes as well as the variations in chromosome meiotic pairings were observed in the reciprocal *Brassica* allopolyploids, but the gene expression changes were not as obvious as the genomic changes (Song et al., 1995; Cui et al., 2012, 2013). Although large-scale differences in gene expression existed between the reciprocal synthetics of *B. napus* (AACC and CCAA), the numbers of genes showing co-up and co-down regulation gave insignificant difference (**Table 4**). It was interesting to see that these co-regulated genes (co-up and co-down) were enriched in the similar functional classes

with similar numbers, revealing that certain gene networks may be particularly susceptible by cytoplasm (**Figure 6**). Overall, the equivalence in either number or function of co-regulated genes indicated that there might be a balance of differentially expressed genes in nuclear-cytoplasmic interactions.

Tissue-specific and Differential Expression of r-protein Genes

Although the ribosome was well-known to be an immensely complex "molecular machine" in translating the genetic code, recent studies indicated that the r-protein genes were likely involved in tissue-specific process and had a regulatory role in the development of plant (Byrne, 2009; Whittle and Krochko, 2009). Our data indicated that those sets of r-protein genes which represented a wide range of r-proteins were differently expressed between tissues (**Supplementary Table 4**). In general, the average expression level of r-protein genes in leaves was higher than that in silique walls (**Supplementary Figure 3**). Nevertheless, an early finding showed that r-protein gene expression was greatest among highly differentiating reproductive tissues (not including silique walls), for example, microspores, embryos, and seeds, using available EST data. However, the lowest levels were also detected in some reproductive tissues, such as anthers, pollen, seed coats (Whittle and Krochko, 2009). Our results demonstrated that the r-protein gene expression levels in silique walls were lower than that in leaves and such difference could be attributable to the fact that silique walls were more mature and the translation activity was not as high as in young leaves. Additionally, in two *Brassica* allotetraploids with B-genome (AABB and BBCC), the expression level of r-protein genes which were highest among samples in leaves turned to be lowest in silique walls. An recent study in the same synthesized allotetraploids showed higher number and percentage of absent bands of B genome than that of A or C genome in AABB and BBCC using AFLP-cDNA, suggesting that gene expression was particularly susceptible to perturbation when B-genome was present in *Brassica* allotetraploids during plant development (Cui et al., 2013).

Hierarchical clustering of these genes indicated that the expression relationships among *Brassica* allotetraploids and diploids were diverse in different tissues (**Figures 7**, **8**). The hierarchy of r-protein genes in leaves (B > A > C) was consistent with that of the expression of rRNA genes (Chen and Pikaard, 1997). However, the expression of rRNA genes could be developmentally regulated and tissue-specific, because the rRNA genes silenced in vegetative tissues were found to be expressed in reproductive tissues, including sepals and petals (Chen and Pikaard, 1997). Thus, the expression divergence of r-protein genes could be associated with the expression change of parental rRNA genes.

In summary, widespread changes of gene expression were observed among the serial resynthesized *Brassica* allotetraploids relative to their diploid progenitors. There were considerable alterations on temporal and spatial expressions, and the range of variations in silique walls was much wider than in leaves. The expression of r-protein genes was tissue-specific and associated with nucleolar dominance. Furthermore, novel gene expression was negatively correlated with silencing during the transition from vegetative to reproductive development. The balance between the transgressive up- and down-expressions was also observed in leaves, as well as between the co-up and co-down regulation in different cytoplasm, such profound changes might enhance the fitness and adaptability of *Brassica* allopolyploids.

ACKNOWLEDGMENTS

This work was funded by National Natural Science Foundation of China (Grant No. 31375656) and by Special Fund for Agro-Scientific Research in the Public Interest (201203026-6).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00957

Supplementary Figure 1 | Analysis of the differentially expressed genes between each allotetraploid and its parents. (A) Leaves. **(B)** Silique walls.

Supplementary Figure 2 | GO functional categories of genes showing novel expression and silencing. (A) Leaves. **(B)** Silique walls.

Supplementary Figure 3 | The average expression level of total r-protein genes among samples.

Supplementary Figure 4 | RT-PCR confirmation of the differentially expressed genes. Columns and bars represent the means and standard error ($n = 3$), respectively. The gene expression levels from RNA-Seq data are added on the top of each gene.

Supplementary Table 1 | Summary of alignment statistics of RNA-Seq in the 14 samples.

Supplementary Table 2 | Top 10 GO items of transgressively regulated genes.

Supplementary Table 3 | Annotation of co-transgressively regulated genes among allotetraploids.

Supplementary Table 4 | R-protein genes matched each *Arabidopsis* r-proteins.

Supplementary Table 5 | The corresponding primers of qRT-PCR.

REFERENCES

Adams, K. L., Cronn, R., Percifield, R., and Wendel, J. F. (2003). Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4649–4654. doi: 10.1073/pnas.0630618100

Adams, K. L., and Wendel, J. F. (2005). Novel patterns of gene expression in polyploid plants. *Trends Genet.* 21, 539–543. doi: 10.1016/j.tig.2005.07.009

Ahuja, I., van Dam, N. M., Winge, P., Traelnes, M., Heydarova, A., Rohloff, J., et al. (2015). Plant defence responses in oilseed rape MINELESS plants after attack by the cabbage moth *Mamestra brassicae*. *J. Exp. Bot.* 66, 579–592. doi: 10.1093/jxb/eru490

Brader, G., Tas, E., and Palva, E. T. (2001). Jasmonate-dependent induction of indole glucosinolates in arabidopsis by culture filtrates of the nonspecific pathogenerwinia carotovora. *Plant Physiol.* 126, 849–860. doi: 10.1104/pp.126.2.849

Buggs, R. J., Doust, A. N., Tate, J. A., Koh, J., Soltis, K., Feltus, F. A., et al. (2009). Gene loss and silencing in *Tragopogon miscellus* (*Asteraceae*): comparison of natural and synthetic allotetraploids. *Heredity (Edinb).* 103, 73–81. doi: 10.1038/hdy.2009.24

Byrne, M. E. (2009). A role for the ribosome in development. *Trends Plant Sci.* 14, 512–519. doi: 10.1016/j.tplants.2009.06.009

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 58, 377–406. doi: 10.1146/annurev.arplant.58.032806.103835

Chen, Z. J., and Pikaard, C. S. (1997).Transcriptional analysis of nucleolar dominance in polyploidy plant: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3442–3447. doi: 10.1073/pnas.94.7.3442

Cui, C., Ge, X., Gautam, M., Kang, L., and Li, Z. (2012). Cytoplasmic and genomic effects on meiotic pairing in *Brassica* hybrids and allotetraploids from pair crosses of three cultivated diploids. *Genetics* 191, 725–738. doi: 10.1534/genetics.112.140780

Cui, C., Ge, X., Zhou, Y., Li, M., and Li, Z. (2013). Cytoplasmic and genomic effects on non-meiosis-driven genetic changes in *Brassica* hybrids and allotetraploids from pairwise crosses of three cultivated diploids. *PLoS ONE* 8:e65078. doi: 10.1371/journal.pone.0065078

Deng, W., Wang, Y., Liu, Z., Cheng, H., and Xue, Y. (2014). Hemi: a toolkit for illustrating heatmaps. *PLoS ONE* 9:e111988. doi: 10.1371/journal.pone.0111988

Dong, Y. Z., Liu, Z. L., Shan, X. H., Qiu, T., He, M. Y., and Liu, B. (2005). Allopolyploidy in wheat induces rapid and heritable alterations in DNA methylation patterns of cellular genes and mobile elements. *Genetika* 41, 1089–1095. doi: 10.1007/s11177-005-0177-7

Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., et al. (2008). Evolutionary genetics of genome merger and doubling in plants. *Annu. Rev. Genet.* 42, 443–461. doi: 10.1146/annurev.genet.42.110807.091524

Flagel, L. E., and Wendel, J. F. (2010). Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *New Phytol.* 186, 184–193. doi: 10.1111/j.1469-8137.2009.03107.x

Gaeta, R. T., Pires, J. C., Iniguez-Luy, F., Leon, E., and Osborn, T. C. (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* 19, 3403–3417. doi: 10.1105/tpc.107.054346

Ge, X. H., Ding, L., and Li, Z. Y. (2013). Nucleolar dominance and different genome behaviors in hybrids and allopolyploids. *Plant Cell Rep.* 32, 1661–1673. doi: 10.1007/s00299-013-1475-5

Hegarty, M. J., Barker, G. L., Brennan, A. C., Edwards, K. J., Abbott, R. J., and Hiscock, S. J. (2008). Changes to gene expression associated with hybrid speciation in plants: further insights from transcriptomic studies in Senecio. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3055–3069. doi: 10.1098/rstb.2008.0080

Hegarty, M. J., Batstone, T., Barker, G. L., Edwards, K. J., Abbott, R. J., and Hiscock, S. J. (2011). Nonadditive changes to cytosine methylation as a consequence of hybridization and genome duplication in *Senecio* (*Asteraceae*). *Mol. Ecol.* 20, 105–113. doi: 10.1111/j.1365-294X.2010.04926.x

Hiruma, K., Onozawa-Komori, M., Takahashi, F., Asakura, M., Bednarek, P., Okuno, T., et al. (2010). Entry mode–dependent function of an indole glucosinolate pathway in *Arabidopsis* for nonhost resistance against anthracnose pathogens. *Plant Cell* 22, 2429–2443. doi: 10.1105/tpc.110.074344

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, 4, 44–57.

Jiang, J., Shao, Y., Du, K., Ran, L., Fang, X., and Wang, Y. (2013). Use of digital gene expression to discriminate gene expression differences in early generations of resynthesized *Brassica napus* and its diploid progenitors. *BMC Genomics* 14:72. doi: 10.1186/1471-2164-14-72

Leitch, A. R., and Leitch, I. J. (2008). Genomic plasticity and the diversity of polyploid plants. *Science* 320, 481–483. doi: 10.1126/science.1153585

Li, A., Liu, D., Wu, J., Zhao, X., Hao, M., Geng, S., et al. (2014). mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* 26, 1878–1900. doi: 10.1105/tpc.114.124388

Liu, S., Liu,Y, Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930. doi: 10.1038/ncomms4930

Nicolas, S. D., Le Mignon, G., Eber, F., Coriton, O., Monod, H., Clouet, V., et al. (2007). Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids. *Genetics* 175, 487–503. doi: 10.1534/genetics.106.062968

Nicolas, S. D., Monod, H., Eber, F., Chèvre, A. M., and Jenczewski, E. (2012). Non-random distribution of extensive chromosome rearrangements in *Brassica napus* depends on genome organization. *Plant J.* 70, 691–703. doi: 10.1111/j.1365-313X.2012.04914.x

Otto, S. P. (2007). The evolutionary consequences of polyploidy. *Cell* 131, 452–462. doi: 10.1016/j.cell.2007.10.022

Prakash, S., Bhat, S., Quiros, C. F., Kirti, P. B., and Chopra, V. L. (2009). Brassica and its close allies: cytogenetics and evolution. *Plant Breed. Rev.* 31, 21–187. doi: 10.1002/9780470593783.ch2

Rapp, R. A., Udall, J. A., and Wendel, J. F. (2009). Genomic expression dominance in allopolyploids. *BMC Biol.* 7:18. doi: 10.1186/1741-7007-7-18

Salmon, A., Ainouche, M. L., and Wendel, J. F. (2005). Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (*Poaceae*). *Mol. Ecol.* 14, 1163–1175. doi: 10.1111/j.1365-294X.2005.02488.x

Soltis, P. S., Liu, X., Marchant, D. B., Visger, C. J., and Soltis, D. E. (2014). Polyploidy and novelty: Gottlieb's legacy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130351. doi: 10.1098/rstb.2013.0351

Soltis, P. S., and Soltis, D. E. (2012). *Polyploidy and Genome Evolution*. Berlin: Springer.

Song, K., Lu, P., Tang, K., and Osborn, T. C. (1995). Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7719–7723. doi: 10.1073/pnas.92.17.7719

Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Coriton, O., Jenczewski, E., et al. (2011). Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. *New Phytol.* 191, 884–894. doi: 10.1111/j.1469-8137.2011.03729.x

Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Huneau, C., Belcram, H., et al. (2010). The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytol.* 186, 102–112. doi: 10.1111/j.1469-8137.2010.03182.x

Wang, J., Tian, L., Lee, H. S., Wei, N. E., Jiang, H., Watson, B., et al. (2006). Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* 172, 507–517. doi: 10.1534/genetics.105.047894

Wang, J., Tian, L., Madlung, A., Lee, H. S., Chen, M., Lee, J. J., et al. (2004). Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* 167, 1961–1973. doi: 10.1534/genetics.104.027896

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Whittle, C. A., and Krochko, J. E. (2009). Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* 21, 2203–2219. doi: 10.1105/tpc.109.068411

Xiong, Z., Gaeta, R. T., and Pires, J. C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7908–7913. doi: 10.1073/pnas.1014138108

Xu, Y., Zhao, Q., Mei, S., and Wang, J. (2012). Genomic and transcriptomic alterations following hybridisation and genome doubling in trigenomic allohexaploid *Brassica carinata* x *Brassica rapa*. *Plant Biol. (Stuttg).* 14, 734–744. doi: 10.1111/j.1438-8677.2011.00553.x

Xu, Y., Zhong, L., Wu, X., Fang, X., and Wang, J. (2009). Rapid alterations of gene expression and cytosine methylation in newly synthesized *Brassica napus* allopolyploids. *Planta* 229, 471–483. doi: 10.1007/s00425-008-0844-8

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., et al. (2006). WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 34, W293–W297. doi: 10.1093/nar/gkl031

Yoo, M. J., Szadkowski, E., and Wendel, J. F. (2013). Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity (Edinb).* 110, 171–180. doi: 10.1038/hdy.2012.94

Zhang, X. (2008). The epigenetic landscape of plants. *Science* 320, 489–492. doi: 10.1126/science.1153996

Zhao, Q., Zou, J., Meng, J., Mei, S., and Wang, J. (2013). Tracing the transcriptomic changes in synthetic Trigenomic allohexaploids of *Brassica* using an RNA-Seq approach. *PLoS ONE* 8:e68883. doi: 10.1371/journal.pone.0068883

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* 39, 61–69. doi: 10.1038/ng1929

Conflict of Interest Statement: The reviewer Genlou Sun declares that, despite having previously collaborated with the authors Xianhong Ge, Yujiao Shao, and Zaiyun Li, the review process was handled objectively. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Zhang, Pan, Cui, Tan, Ge, Shao and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 October 2015
doi: 10.3389/fpls.2015.00894

The miRNAs and their regulatory networks responsible for pollen abortion in Ogura-CMS Chinese cabbage revealed by high-throughput sequencing of miRNAs, degradomes, and transcriptomes

Xiaochun Wei[1†], Xiaohui Zhang[2†], Qiuju Yao[1], Yuxiang Yuan[1], Xixiang Li[2], Fang Wei[3], Yanyan Zhao[1], Qiang Zhang[1], Zhiyong Wang[1], Wusheng Jiang[1] and Xiaowei Zhang[1]*

[1] Institute of Horticulture, Henan Academy of Agricultural Sciences, Zhengzhou, China, [2] Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China, [3] College of Life Science, Zhengzhou University, Zhengzhou, China

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Maoteng Li,
Huazhong University of Science and Technology, China
Hao Peng,
Washington State University, USA

***Correspondence:**
Xiaowei Zhang
xiaowei5737@163.com

[†]These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: 29 July 2015
Accepted: 08 October 2015
Published: 21 October 2015

Citation:
Wei X, Zhang X, Yao Q, Yuan Y, Li X, Wei F, Zhao Y, Zhang Q, Wang Z, Jiang W and Zhang X (2015) The miRNAs and their regulatory networks responsible for pollen abortion in Ogura-CMS Chinese cabbage revealed by high-throughput sequencing of miRNAs, degradomes, and transcriptomes. Front. Plant Sci. 6:894. doi: 10.3389/fpls.2015.00894

Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) is one of the most important vegetables in Asia and is cultivated across the world. Ogura-type cytoplasmic male sterility (Ogura-CMS) has been widely used in the hybrid breeding industry for Chinese cabbage and many other cruciferous vegetables. Although, the cause of Ogura-CMS has been localized to the orf138 locus in the mitochondrial genome, however, the mechanism by which nuclear genes respond to the mutation of the mitochondrial orf138 locus is unclear. In this study, a series of whole genome small RNA, degradome and transcriptome analyses were performed on both Ogura-CMS and its maintainer Chinese cabbage buds using deep sequencing technology. A total of 289 known miRNAs derived from 69 families (including 23 new families first reported in *B. rapa*) and 426 novel miRNAs were identified. Among these novel miRNAs, both 3-p and 5-p miRNAs were detected on the hairpin arms of 138 precursors. Ten known and 49 novel miRNAs were down-regulated, while one known and 27 novel miRNAs were up-regulated in Ogura-CMS buds compared to the fertile plants. Using degradome analysis, a total of 376 mRNAs were identified as targets of 30 known miRNA families and 100 novel miRNAs. A large fraction of the targets were annotated as reproductive development related. Our transcriptome profiling revealed that the expression of the targets was finely tuned by the miRNAs. Two novel miRNAs were identified that were specifically highly expressed in Ogura-CMS buds and sufficiently suppressed two pollen development essential genes: *sucrose transporter SUC1* and *H^+-ATPase 6*. These findings provide clues for the contribution of a potential miRNA regulatory network to bud development and pollen engenderation. This study contributes new insights to the communication between the mitochondria and chromosome and takes one step toward filling the gap in the regulatory network from the orf138 locus to pollen abortion in Ogura-CMS plants from a miRNA perspective.

Keywords: miRNAs, *Brassica rapa* ssp. *pekinensis*, Ogura-CMS, bud, pollen, deep sequencing

INTRODUCTION

Cytoplasmic male sterility (CMS) is a maternally inherited trait that results in the loss of the ability to produce fertile pollen; CMS has been used extensively for hybrid crop breeding (Woodson and Chory, 2008; Luo et al., 2013). Ogura-CMS was originally discovered in the wild radish (*Raphanus sativus*) (Ogura, 1968) and has been widely introduced into Cruciferous vegetables and oil crops and successfully applied in the heterosis breeding industry due to its complete male sterility and stability (Pelletier et al., 1983). Ogura-CMS is controlled by a mitochondrial orf138 locus that consists of two co-transcribed open reading frames: orf138 and atp8 (ATP synthase, subunit 8) (Krishnasamy and Makaroff, 1993; Grelon et al., 1994). However, the mechanism by which the orf138 locus results in male sterility is unclear. Many researchers have focused on the crosstalk between the mitochondria and nucleus, and the CMS is believed to be controlled by mitochondrial-nuclear interactions (Chase, 2007). Several CMS-associated nuclear genes, such as those related to programmed cell death (PCD) and reactive oxygen species (ROS), have been identified and proposed to play roles in the CMS pathway (Balk and Leaver, 2001). Other regulatory signaling pathways, such as the biogenesis of jasmonic acid, are also impaired in the CMS line (Liu et al., 2012).

MicroRNAs (miRNAs) are a class of 21–24 nt small non-coding RNAs that regulate gene expression by post-transcriptional repression (Carrington and Ambros, 2003; Bartel, 2004). In plants, miRNAs are generated from primary miRNA transcripts (pri-miRNAs) that are generally transcribed by RNA polymerase II (Bartel, 2004; Khraiwesh et al., 2010). Primary transcripts containing a distinctive hairpin structure are first trimmed by DICER-LIKE1 (DCL1) to generate miRNA precursors (pre-miRNAs) in the nucleus. Then, these pre-miRNAs are transported to the cytoplasm and further processed by DCL1 to generate ∼21 nt mature miRNAs (Kurihara and Watanabe, 2004). Subsequently, the mature miRNAs are loaded onto the RNA-induced silencing complex (RISC) and guide RISC recognition of complementary sites on target mRNAs, thereby inducing transcript cleavage (Bartel, 2004; Baulcombe, 2004) or translational repression (Aukerman and Sakai, 2003; Chen, 2004). Plant miRNAs mainly function by directing the cleavage of their highly complementary target transcripts. Thus, it is easy to predict and verify their targets using bioinformatics and experimental methods.

miRNAs play important roles in the regulation of a wide range of plant developmental processes, including plant architecture (Zhang et al., 2011b), leaf development (Floyd and Bowman, 2004; Juarez et al., 2004), root development (Guo et al., 2005), tuberization (Bhogale et al., 2014), the vegetative-reproductive phase change (Wang et al., 2011a), flowering time (Zhou et al., 2013), floral organ identity (Aukerman and Sakai, 2003; Bartel, 2004; Chen, 2004; Nagpal et al., 2005), self-incompatibility (Tarutani et al., 2010), plant nutrient homeostasis (Yamasaki et al., 2007; He et al., 2014), and response to environmental biotic and abiotic stresses (Juarez et al., 2004; Navarro et al., 2006; Sunkar et al., 2007; Katiyar-Agarwal and Jin, 2010; Zhang et al., 2011a). However, few miRNAs have been confirmed to control pollen development or CMS. Efforts have been made to screen for pollen-specific or enriched miRNAs in rice (Wei et al., 2011), Arabidopsis (Chambers and Shuai, 2009; Grant-Downton et al., 2009a,b), and pakchoi (Jiang et al., 2014).

Chinese cabbage (*Brassica rapa* ssp. *pekinensis*) is an important leafy vegetable that is a cross-pollinated crop with significant heterosis. Ogura-CMS was transferred to the Chinese cabbage in the 1980s and is still widely used in the breeding industry (Yamagishi and Bhat, 2014). The miRNAs from the Ogura-CMS and the maintainer line buds of a closely related subspecies (*B. rapa* ssp. *chinensis*) have been profiled (Jiang et al., 2014). In that study, 54 new conserved miRNAs and 25 pairs of novel miRNAs were identified. Among them, 18 miRNAs were differentially expressed between the male sterile and fertile lines. A genome-wide analysis of miRNAs in *B. rapa* by deep sequencing has been reported (Kim et al., 2012). However, to the best of our knowledge the miRNAs that underlie flower bud development and respond to the Ogura-CMS in Chinese cabbage have not been examined.

In present study, the miRNAs from the buds of a Chinese cabbage Ogura-CMS line (Tyms) and its maintainer line (231-330) were profiled by small RNA deep sequencing. The targets of the miRNAs were identified by degradome sequencing analysis. The expression levels of the corresponding targets were monitored by transcriptome sequencing. These results provide a full set of miRNA-target messages that underlie flower bud development in the Chinese cabbage and provide insights into miRNA-related regulatory functions that may play roles in the mitochondrial–nuclear interactions in Ogura-CMS.

MATERIALS AND METHODS

Plant Materials and RNA Extraction

The Chinese cabbage Ogura-CMS sterile line (Tyms) and its maintainer line (231-330) used in this study were grown in the Henan Academy of Agricultural Sciences, Yuanyang, Henan Province, China. The lines possessed isogenic chromosomes with different cytoplasmic genes. Flower buds <6 mm length were stripped from 10 individual of each genotype of plants. These buds contain the entire developmental progress from anther generation to pollen abortion. Tyms and 231-330 were sampled separately and then snap-frozen in liquid nitrogen and kept at -80°C for further use. Total RNA was extracted using the TRIzol reagent (Invitrogen, USA). DNase (Promega, USA) was used to remove potential DNA contamination.

Small RNA Library Construction and Sequencing

The RNA samples from 231-330 (Control) and Tyms (CMS) were quantified and equalized so that equivalent amounts of RNA were analyzed. A total of 30 μg of RNA was resolved on denatured polyacrylamide gels. Gel fragments with the size range of 18–30 nt were excised and recovered. These small RNAs were ligated with 5′ and 3′RNA adapters with the T4 RNA ligase. The adapter-ligated small RNAs were subsequently transcribed into cDNA by Super-Script II Reverse Transcriptase (Invitrogen) and amplified

using primers specific for the ends of the adapters. The amplified cDNA products were purified and finally sequenced using Solexa sequencing technology (BGI, Shenzhen, China).

Identification of Known and Novel miRNAs

The adapter sequences, impurities, and sequences with less or more than 18–30 nt were filtered out from the raw sequence reads. The remaining sequences that ranged from 18 to 30 nt in length were used for known miRNA prediction. First, we aligned the tags to the *B. rapa* miRNA precursors in miRBase (version 20.0, http://www.mirbase.org/index.shtml) with no mismatches allowed. Then, the obtained tags were aligned to the mature miRNAs of *B. rapa* with at least a 16 nt overlap to allow offsets. The miRNAs that satisfied both of the above criteria were counted to obtain the expression number of identified miRNAs. The rest of the small RNA tags were aligned to the miRNA precursors/mature miRNAs of all plants in miRBase, allowing two mismatches and free gaps. The miRNAs with the highest expression levels for each mature miRNA family were chosen as a temporary miRNA database. Then, the precursors of the temporary miRNAs in the *B. rapa* genome were predicted. Those that failed to fold into hairpin structures were regarded as pseudo-miRNAs and discarded, while those that fulfilled the miRNA criteria were adopted as newly identified known miRNAs.

By comparing our sequences with those in the databases and the Chinese cabbage genome, the sRNAs can be annotated into different categories, including siRNA, piRNA, rRNA, tRNA, snRNA, snoRNA, repeat associated sRNA, degraded tags of exons or introns, and sRNAs that could not be annotated. The tags annotated as intron, exon antisense, and unknown were used to predict novel miRNAs using the software Mireap. The key conditions are as follow: hairpin miRNAs can fold into secondary structures and mature miRNAs are present in one arm of the hairpin precursors; the 5-p and 3-p mature miRNAs present 2-nucleotide 3′ overhangs; hairpin precursors lack large internal loops or bulges; the secondary structures of the hairpins are steady, with the free energy of hybridization lower than or equal to −18 kcal/mol; and the copy number of mature miRNAs with predicted hairpins must be greater than five in the alignment result. The expression of novel miRNAs was produced by summing the count of miRNAs with no more than three mismatches on the 5′ and 3′ ends and no mismatches in the middle from the alignment result.

The differentially expressed miRNAs were calculated with the following procedures. First, the expression of miRNAs in two samples was normalized to obtain the expression of transcript per million (TPM). The normalization formula was as follows: Normalized expression = Actual miRNA count/Total count of clean reads × 1,000,000; Second, the fold-change and *P*-value were calculated from the normalized expression.

Degradome Library Construction and Target Identification

Total RNA extracted from the 231-330 (Control) and Tyms (CMS) lines. Approximately 200 μg of the total RNA was polyadenylated using the Oligotex mRNA mini kit (Qiagen). A 5′ RNA adapter was added to the cleavage products (which possessed a free 5′-monophosphate at their 3′ termini) using the T4 RNA ligase (Takara). Then, the ligated products were purified using the Oligotex mRNA mini kit (Qiagen) for reverse transcription to generate the first strand of cDNA using an oligo dT primer via SuperScript II RT (Invitrogen). After the cDNA library was amplified for 6 cycles (94°C for 30 s, 60°C for 20 s, and 72°C for 3 min) using Phusion Taq (NEB), the PCR products were digested with the restriction enzyme Mme I (NEB). A double-stranded DNA adapter was ligated to the digested products using T4 DNA ligase (NEB). The ligated products were selected based on size in a 10% polyacrylamide gel and purified for the final PCR amplification (94°C for 30 s, 60°C for 20 s, and 72°C for 20 s) for 20 cycles (German et al., 2009). The PCR products were gel purified and used for high-throughput sequencing with the Illumina HiSeq 2000.

Low quality sequences and adapters were removed, and the unique sequence signatures were aligned to the database of Chinese cabbage transcript assemblies in the Brassica Gene Index (http://brassicadb.org/brad/downloadOverview.php) using the SOAP software (Li et al., 2008) (http://soap.genomics.org.cn/). CleaveLand was used to detect potentially cleaved targets based on degradome sequences. The 20 and 21 nt distinct reads were subjected to the CleaveL and pipeline for small RNA target identification as previously described (Addo-Quaye et al., 2009). The tags mapped to cDNA sense strands were used to predict cleavage sites. The miRNA-mRNA pairs were searched and *p*-values were calculated using PAREsnip (http://srna-workbench.cmp.uea.ac.uk/tools/paresnip/). Only *p*-values less than 0.05 were adopted for the *t*-plot figure (Folkes et al., 2012). All alignments with scores not exceeding 4 that possessed the 5′ end of the degradome sequence coincident with the 10th and 11th nucleotides that were complementary to the small RNA were retained. To evaluate the potential functions of miRNA-targeted genes, gene ontology (GO) categories (http://www.geneontology.org/) were used to assign the identified target genes according to the previously described method (Du et al., 2010).

Transcriptome Sequencing and Target Gene Profiling

Total RNA (10 μg) was subjected to poly-A selection, fragmentation, random priming, first and second strand cDNA synthesis with the Illumina Gene Expression Sample Prep kit (CA, USA). The cDNA fragments were subjected to an end repair process and then ligated to adapters. The products were enriched with PCR, and the 200-bp fragments were purified with 6% TBE PAGE gel electrophoresis. After denaturation, the single-chain fragments were fixed onto the Solexa Sequencing Chip (Flowcell) and consequently grown into single-molecule cluster sequencing templates through in situ amplification on the Illumina Cluster Station. Double-end pyrosequencing was performed on the Illumina Genome Analyzer platform with read lengths of 101 bp for each end. The clean reads were aligned to the Chinese cabbage reference transcript assemblies (http://brassicadb.org/brad/downloadOverview.php). Gene expression levels were calculated using the RPKM method (Mortazavi et al., 2008) using

the following formula: RPKM = (1,000,000 * C)/(N * L * 1000), where RPKM(A) is the expression of gene A, C is the number of reads that uniquely align to gene A, N is the total number of reads that uniquely align to all genes, and L is the number of bases in gene A. Statistical comparison between Tyms and 231–330 was performed using the IDEG6 software (Romualdi et al., 2003). The General Chi-squared method was used and the FDR (false discovery rate) was applied to determine the Q-value threshold. Unigenes were considered to be differentially expressed when the RPKM between Tyms and 231–330 displayed a more than two-fold change with an FDR less than 10^{-2}.

Quantitative Real-time PCR

For analysis of miRNAs, 2.5 μg of total RNA was polyadenylated using a miRNA cDNA synthesis kit (Takara, Inc., Dalian, China). The poly(A)-tail-amended total RNA was reverse-transcribed by PrimeScript RTase using a universal adapter primer containing oligo-dT. The qPCR was performed on a LightCycler 96 System (ROCHE, USA) using SYBR Premix Ex TaqTMII (TaKaRa, Dalian China). The miRNA-specific forward primer for each miRNA was designed based on the entire miRNA sequence (Table S13), and the universal reverse primer was provided by the miRNA cDNA synthesis kit (Takara, Dalian, China). All reactions were performed with three biological and three technical replicates for each sample, and the U6 snRNA (Forward: GGGGACATCCGATAAAATT, Reverse: TGTGCGTGTCATCCTTGC) was used as the internal control. The reaction volume was 20 μL, including 10 μL of SYBR *Premix Ex Taq* II, 0.8 μL of 10 mM Forward primer, 0.8 μL of 10 mM Uni-miR qPCR primer, 2.0 μL of the cDNA sample and 6.4 μL of dH_2O. The following qPCR program was used: denaturation at 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 55°C for 30 s, and 72°C for 60 s. Melting curve analysis with 95°C for 10 s, 65°C for 60 s, and 97°C for 1 s was performed to produce a dissociation curve for verification of the amplification specificity. Relative expression levels of miRNAs were quantified using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

The primers of the selected genes subjected to target analysis are listed in Table S14. β-*actin* was used as an internal control. Experiments were performed on a similar system as described above. The reaction volume was 20 μL, including 10 μL of SYBR® *Premix Ex Taq*™ (Tli RNaseH Plus), 0.8 μL of 10 mM Forward primer, 0.8 μL of 10 mM Reverse primer, 2.0 μL of the cDNA sample and 6.4 μL of dH_2O. Three independent biological and three technical replicates were performed. The fold change was estimated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Paraffin Sectioning and Microscopic Observation

Sterile and fertile flower buds were fixed and embedded in Paraffin (Beeswax, China). Thin (0.8 μm) sections were prepared with an Ultracut Eultra microtome (Leica, Germany), stained with hematoxylin, and photographed under a LEICA DMI3000B microscope (Leica, Germany).

RESULTS AND DISCUSSION

sRNA Sequencing and miRNA Identification

To analyze the roles of miRNAs in Ogura-CMS in Chinese cabbage, small RNAs were pyrosequenced from the buds of the Ogura-CMS Tyms line and its maintainer line 231–330 (the morphology and microscopy of the samples are shown in **Figure 1**). A total of 22.2 and 24.4 M reads were produced from the two lines, respectively. After filtering out the low quality reads, 3′adapter null, insert null, 5′adapter contaminants, reads smaller than 18 nt and poly A, a total of 21.9 and 24.1 M high quality clean reads were obtained, respectively (Table S1). The majority of the tags ranged in size between 21–24 nt, with the 24 and 21 nt lengths dominant (**Figure 2**). Out of the 46.0 M total tags, 33.8 M (73.5%) were shared by the two samples, while 5.7 M (12.5%) and 6.4 M (14.0%) were Tyms- and 231–330-specific, respectively. Then, the tags were divided into 12.2 M non-redundant unique reads composed of 4.8 M (39.6%) Tyms-specific, 5.6 M (45.7%) 231–330-specific and 1.8 M (14.8%) reads shared by both lines (Table S2).The reads were aligned to *B. rapa* miRNAs (*B. rapa* 1.1) in miRBase (http://www.mirbase.org/); among the 157 mature miRNAs encoded by 96 precursors distributed into 63 families included in the database, 69 mature miRNAs from 59 precursors of 41 families were expressed in our samples (Table S3). The remainder of the reads were aligned to all known plant miRNAs in miRBase 21.0 and then aligned to the *B. rapa* genome for precursor identification. From this analysis, another 220 mature miRNAs derived from 163 precursors were identified. Among them, 57 precursors contained both 5p and 3p miRNAs, while only one mature miRNA was detected for the remaining 106 precursors. The 139 (39 pairs and 61 singles) mature miRNAs encoded by 100 precursors were new members of 36 existing families in the *B. rapa* miRBase. Thus, only 17 families (bra-MIR9552—bra-MIR9557 and bra-MIR9559—bra-MIR9569) in the miRBase *B. rapa* collection were not detected in our present study. By expending the members to the existing *B. rapa* miRNA families, bra-miR156 was now the largest family (harboring 21 members), followed by bra-miR171 (11 members), bra-miR167 (8 members), bra-miR172 (8 members), bra-miR164 (7 members), bra-miR168 (7 members), bra-miR2111 (7 members), bra-miR157 (6 members), bra-miR160 (6 members), bra-miR390 (6 members), and bra-miR395 (6 members). The other families contained less than five members (Table S4). In addition to these existing families, we also identified 81 miRNAs (18 pairs and 45 singles from 63 precursors) belonging to 23 families that had not been previously reported in *B.rapa*. These included 13 members of bra-miR1439, seven members of bra-miR169, five members each of bra-miR166, bra-miR393, bra-miR394, and bra-miR399, three bra-miR165, two bra-miR170, two bra-miR5376, and one member each of bra-miR397, bra-miR827, bra-miR828, bra-miR838, bra-miR845, bra-miR858, bra-miR5298, bra-miR5575, bra-miR5641, bra-miR6029, bra-miR6030, bra-miR6033, bra-miR6034, and bra-miR6284 (Table S5). The remaining reads

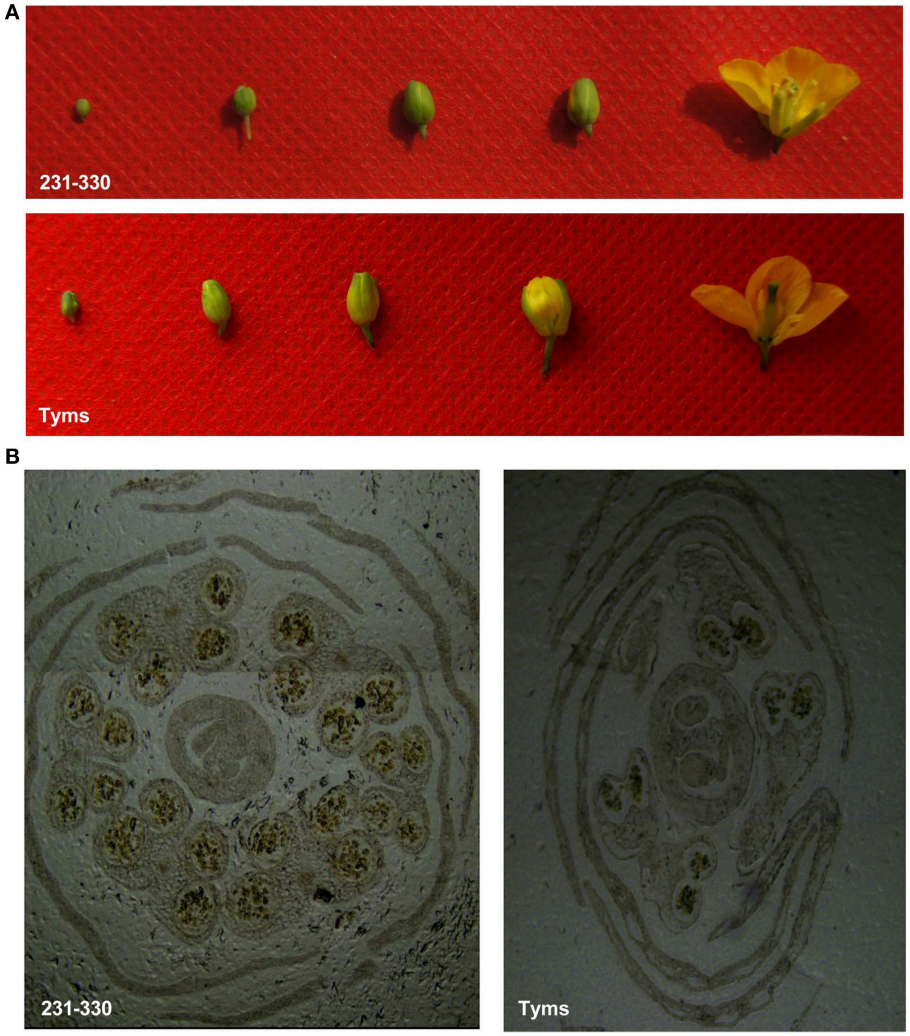


FIGURE 1 | The buds and flowers of the Ogura-CMS (Tyms) and maintainer (231-330) lines. (A) the buds of four developmental stages and the flower that were sampled for this study. **(B)** microscopy of the buds showing the well-developed pollen sacs in 231–330 and incomplete development in Tyms.

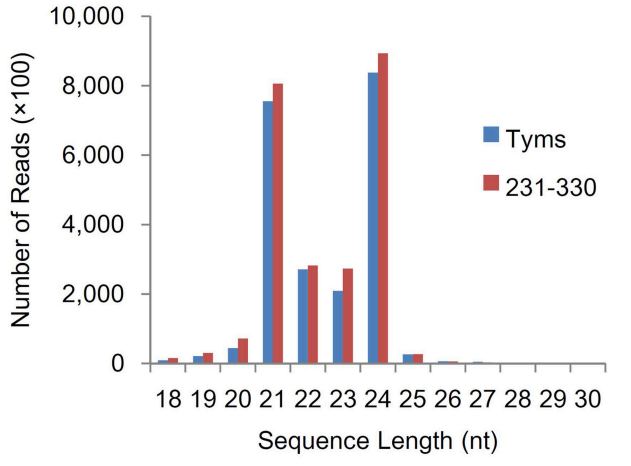


FIGURE 2 | Size distribution of small RNA sequences.

were aligned to Genbank, Rfam and the cabbage genome and annotated as rRNA, scRNA, snRNA, snoRNA, tRNA, repeat, exon sense, exon antisense, intron sense, intron antisense, and unannotated sequences (Table S6). The unannotated reads and those derived from the intron region and exon antisense region were used for novel miRNA analysis. A total of 426 novel miRNAs were identified based on the criteria defined in the Methods section. The sequences and precursor information were listed in Table S7. Among them, 83 novel miRNAs were generated from more than one precursor in the Chinese cabbage. A total of 138 precursors (84 types of mature miRNA) harbored both the 3-p and 5-p miRNAs on their arms (Supplementary file 2).

Of these miRNAs, members of 32 known families were also identified in buds of a closely related vegetable plant *B. rapa* ssp. *chinensis* (Jiang et al., 2014). These miRNAs may be common bud development contributors in both plants. Our present study identified more miRNAs of either known or novel

type than that of the previous report. One important reason for this is that nearly fourfold of reads were generated in present study compared to the previous one. This facilitated the identification of low abundant miRNAs. Another reason can be attributed to the genetic differences between the two subspecies. One of our sequencing projects showed only half of reads from *B. rapa* ssp. *chinensis* could map to the *B. rapa* ssp. *pekinensis* reference genome (unpublished data). The publicly released *B. rapa* ssp. *pekinensis* genome (Wang et al., 2011b) is not an ideal reference for *B. rapa* ssp. *chinensis*. Jiang et al. (2014) used *B. rapa* ssp. *pekinensis* reference genome for their miRNA annotation and novel-miRNA prediction, there for many miRNAs were failed to be identified due to the sequence diversity, especially for the novel miRNAs which rely on the hairpins prediction.

Comparative Expression Patterns of miRNAs in the Buds of Ogura-CMS and its Maintainer

The expression levels of known and novel miRNAs were profiled based on tag counts. For multiple precursors sharing the same mature miRNA, only one member of the mature miRNA was used for expression profiling. The miRNA with the highest expression levels was bra-miR157, which accumulated 579,691 and 253,845 copies in the 231-330 and Tyms buds, respectively, followed by bra-miR168 and bra-miR156 (Table S3). bra-miR156 and bra-miR157 belong to the same family and regulate the *SQUAMOSA promoter binding protein-like* (*SPL*) genes that are important regulators for plant vegetative and reproductive development as well as gynoecium differential patterning and male fertility (Wang et al., 2009; Wu et al., 2009; Xing et al., 2010, 2013; Zhang et al., 2011b; Yu et al., 2012). bra-miR168 targets *ARGONAUTE 1* (*AGO1*), which is an important part of the RISC. The balance between miR168 and AGO1 plays important roles in plant development, including flowering and fruiting (Vaucheret et al., 2004; Xian et al., 2014). The top 20 most highly expressed known miRNAs are listed in **Table 1**. The expression levels of novel miRNAs were much lower than the known miRNAs, with the majority (∼80%) barely accumulating several to less than 100 copies. The top 15 most highly expressed novel miRNAs are listed in **Table 2**.

The expression level between the Ogura-CMS and maintainer buds was compared using normalized tag counts. Ten known miRNAs were down-regulated two-fold ($p < 10^{-3}$) in Tyms compared to 231–330 (**Table 3**). Among them, bra-miR157, bra-miR158-3p, and bra-miR5718 were expressed at relatively high levels and could play important roles in pollen development. However, bra-miR6030 was the only known miRNA that was up-regulated in the Tyms buds. In addition to these examples, a total of 49 and 27 novel miRNAs were down- and up-regulated in Tyms, accounting for 11.5 and 6.3% of the total 426 novel miRNAs, respectively (Table S8).

To test the accuracy of the RNA-Seq-based expressional profiling, a set of Q-PCR analyses were performed. As shown in **Figure 3**, the relative expression levels were similar between the Q-PCR and RNA-Seq technologies for seven known and one novel miRNAs. These results indicate that the RNA-Seq expression profiles are reliable.

TABLE 1 | The most highly expressed known miRNAs.

miRNA	Count		Sequence
	231-330	Tyms	
bra-miR157	579,691	253,845	TTGACAGAAGATAGAGAGCAC
bra-miR168	506,969	373,990	TCGCTTGGTGCAGGTCGGGAC
bra-miR156	254,093	231,666	TTGACAGAAGAAAGAGAGCAC
bra-miR166	151,210	170,161	TCGGACCAGGCTTCATTCCCC
bra-miR158-5p	66,094	51,736	CTTTGTCTATCGTTTGGAAAAG
bra-miR158-3p	44,804	34,621	TTTCCAAATGTAGACAAAGCA
bra-miR5718	39,028	17,531	TCAGAACCAAACACAGAACAAG
bra-miR408-5p	34,067	32,148	ACAGGGAACAAGCAGAGCATG
bra-miR172b-3p	23,598	16,041	AGAATCTTGATGATGCTGCAT
bra-miR390-5p	21,103	15,941	AAGCTCAGGAGGGATAGCGCC
bra-miR164	15,166	12,237	TGGAGAAGCAGGGCACGTGCA
bra-miR165	14,685	15,120	TCGGACCAGGCTTCATCCCCC
bra-miR6034	12,828	12,514	TCTGATGTATATAGCTTTGGG
bra-miR168-3p	9015	5512	CCCGCCTTGCATCAACTGAAT
bra-miR1885	7209	8016	TACATCTTCTCCGCGGAAGCTC
bra-miR6029	6847	4093	TGGGGTTGTGATTTCAGGCTT
bra-miR160-3p	5180	4304	GCGTATGAGGAGCCATGCATA
bra-miR5654	4807	3418	ATAAATCCCAAGCATCATCCA
bra-MIR5721	4345	3606	AAAATGGAGTGGGAAATGGAG
bra-miR393-3p	4325	4780	ATCATGCGATCTCTTCGGATT

TABLE 2 | The most highly expressed novel miRNAs.

miRNA	Sequence	Count		
		231-330	Tyms	Total
novel-miR-154	GGAATGTTGTCTGGCTCGAGG	17,293	17,185	34,478
novel-miR-113	GTCTGGGTGGTGTAGTCGGTT	3337	1862	5199
novel-miR-269	GGACTGTTGTCTGGCTCGAGG	2362	1902	4264
novel-miR-3	GCGTACAGAGTAGTCAAGCATA	1790	1949	3739
novel-miR-310	TTAGCGGAATATAAGAATCGGTT	2398	1127	3525
novel-miR-228	CTTGCATATCTTAGGAGCTTT	1690	1582	3272
novel-miR-74	TTGCTATAGATGGTTTCTGCT	1781	1340	3121
novel-miR-287	AGATCATCCTGCGGCTTCATT	1152	1336	2488
novel-miR-275	TGAAGTGGAGTAGAGTATAATG	1424	730	2154
novel-miR-121	GGAATGTTGTTTGGCTCGAAG	984	887	1871
novel-miR-199	TGGATGATGCTTGGCTCGAGA	446	1160	1606
novel-miR-320	TGGTAGAACGACCGCATGATC	789	756	1545
novel-miR-50	GCAGCACCATCAAGATTCACA	913	507	1420
novel-miR-261	TCTGCTATCAACTTGTAGAGTC	795	612	1407
novel-miR-55	GCAGCATCATCAAGATTCACA	536	607	1143

Target Prediction and Validation by Degradome Analysis

miRNAs function by regulating their target genes and especially by degrading their target mRNAs in plants. Thus, we performed a degradome sequencing analysis to validate the miRNA

TABLE 3 | Differentially expressed known miRNAs.

Pairwise	miR-name	231-330-std	Tyms-std	Fold-change(log2 Tyms/231-330)	*p*-value	Change
231–330–Tyms	bra–miR157	24058.52	11589.58	−1.05	0	Down
231–330–Tyms	bra–miR158–3p	494.17	79.76	−2.63	0	Down
231–330–Tyms	bra–miR169	58.73	27.53	−1.09	3.62E−59	Down
231–330–Tyms	bra–miR394	20.00	9.63	−1.05	4.01E−20	Down
231–330–Tyms	bra–miR827	19.76	8.67	−1.19	9.70E−24	Down
231–330–Tyms	bra–miR845	22.33	11.09	−1.01	7.63E−21	Down
231–330–Tyms	bra–miR858	11.91	5.07	−1.23	9.91E−16	Down
231–330–Tyms	bra–miR5714	1.83	0.82	−1.15	3.21E−3	Down
231–330–Tyms	bra–miR5716	3.44	1.37	−1.33	5.17E−06	Down
231–330–Tyms	bra–miR5718	1619.75	800.40	−1.02	0	Down
231–330–Tyms	bra–miR6030	7.64	20.55	1.43	1.11E−32	Up

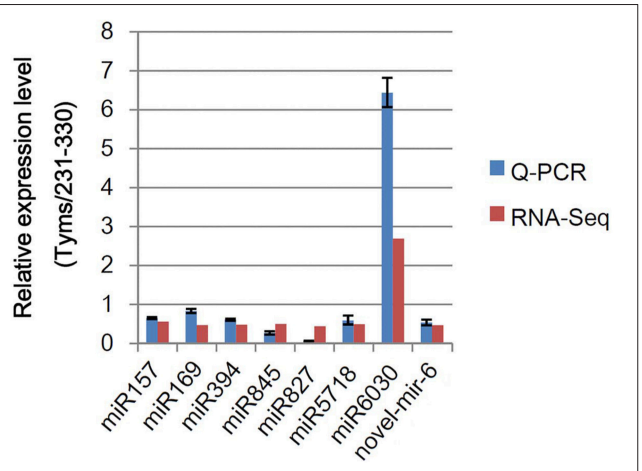


FIGURE 3 | Q-PCR analysis of the relative expression levels of miRNAs. The bars indicate the SE; four biological and three technique replicates were performed.

targets. A total of 36,541,512 and 36,595,548 clean tags were generated from the buds of Tyms and 231-330, respectively (Table S9). Among them, 23,253,666 (63.64%) and 24,768,149 (67.68%) tags were mapped to the reference *B. rapa* genome. A total of 15,154,675 (41.47%) and 17,009,287 (46.48%) tags were mapped to the cDNA sense chains and were used for target prediction (Table S10). A total of 376 mRNAs were identified as miRNA targets, including 136 targeted by 30 known miRNA families and 248 targeted by 100 novel miRNAs. Among which 26 mRNAs were targeted by more than one miRNA (Table S11). A larger proportion known (50.8%) miRNAs compared to novel (23.5%) miRNAs have targets been detected, indicating that the conserved miRNAs have comparatively more valid cutting functions than the newly formed miRNAs. One possible reason for not detecting the targets could be that the target mRNAs were expressed at low levels; thus, the cut ends were not detected by our pipeline. Another possibility is that some miRNAs function by translation inhibition rather than mRNA digestion. Based on GO annotation, the targets were enriched in the "binding," "catalytic activity," and "nucleic acid binding transcription factor activity" terms in the "molecular function" cluster. "Cell," "membrane," and "organelle" were the three most abundant "cellular component" targets. In the "biological process" cluster, the top three terms were "cellular process," "metabolic process," and "response to stimulus" (**Figure 4**).

Most of the known miRNAs targeted transcription factor-encoding genes, such as the targeting of *SPL, MYB, auxin response factor (ARF), NAC, scarecrow, APETALA 2 (AP2), GROWTH-REGULATING FACTOR (GR)*, and *C3HC4-type RING finger* family transcription factors by miR156/157, miR159, miR160, miR164, miR171, miR172, miR396, and miR5716, respectively. This finding indicates that the transcription factors perform primary important roles in bud development. Another large group of miRNAs, including miR158, miR161, miR400, and miR5654, target the pentatricopeptide repeat (PPR)-containing protein-encoding genes. PPR-containing proteins are a large family in plants that are mostly located in the mitochondria and chloroplast and play important roles in RNA processing in the two organelles (Fujii and Small, 2011). Some PPR-containing proteins have been identified as fertility restoration (Rf) genes for CMS (Desloire et al., 2003; Wang et al., 2008; Yasumoto et al., 2009). miR1885, miR5719, and miR6030 target disease resistance (CC-NBS-LRR class) genes. miR167 targets *IAA-amino acid hydrolase 3* (*IAR3*), which encodes an auxin conjugate hydrolase that contributes to the jasmonate pathway (Kinoshita et al., 2012; Widemann et al., 2013). MiR393 targets *AUXIN signaling F-BOX* (*AFB*) genes, which are auxin receptors that have been widely identified as stress response regulators and control auxin-related development in plants (Si-Ammour et al., 2011; Terrile et al., 2012). These genes and the above mentioned miR160-ARFs indicate that the miRNA-hormone signaling cascades play important roles in flower bud development. Interestingly, three *Vacuolar ATP synthase subunit A* (*VHA-A*) tags were identified as targets of miR5712. *VHA-A* have been demonstrated to be essential for male gametophyte development due to its important role in Golgi organization

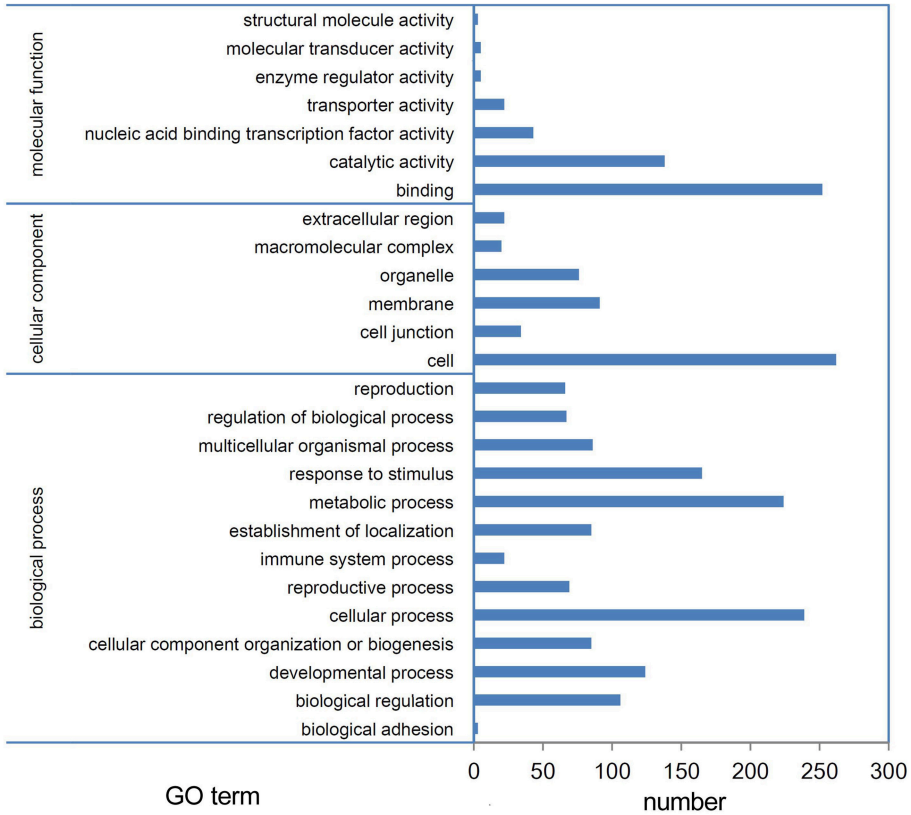


FIGURE 4 | GO annotation of targets. Targets of miRNAs expressed in Tyms and 231-330 Chinese cabbage buds were detected by degradome sequencing analysis.

(Dettmer et al., 2005). miR5724 targets *embryo defective 1473* (*emb1473*), which encodes a structural constituent of the ribosome. miR158, which is down-regulated in Tyms buds, targets a glutathione S-transferase gene that has been annotated as a restorer-of-fertility (Krishnasamy and Makaroff, 1994; Woo et al., 2008). These miRNAs could play important roles in pollen development.

For the novel miRNAs, 153 out of the 264 targets (58%) were annotated as "catalytic activity," which accounted for the largest fraction of the targets. Only 16 transcription factors were identified as targets of novel miRNAs: four *NAC* (including three *cup-shaped cotyledon*) targeted by novel mir410, 446, and 465, four zinc finger family proteins, two homeobox proteins, one *MYB*, one *WRKY*, one *SPL*, one *HSFA1E*, and one *PIL1* (*Phytochrome Interacting Factor 3-like 1*). Several of the 19 targets possess "transporter activity," including four genes encoding H^+-transporting ATPases, two genes encoding Ca^{2+}:H^+ antiporters and two genes encoding sucrose transporter SUC1. These genes could play important roles in gametophyte development by transporting protons, ions, mineral elements and energy. A total of 28 targets were annotated as reproductive process-related, including the aforementioned *cup-shaped cotyledon 2* transcription factors and *sucrose transporter SUC1*, *auxin-responsive protein*, and *ubiquitin-activating enzyme E1* (Table S12). Among them, an oxidoreductase, a HSFA1E, a C2H2 zinc finger protein 1 and a beta-glucosidase were targets of novel-miR-6, −93, and −321, which were down-regulated in Tyms. An F-box/LRR-repeat protein, a heat shock 70 kDa protein 1/8, an auxin-responsive protein IAA and two sucrose transporter SUC1 were targeted by novel-miR-175, −383, −385, and −448, which were up-regulated in Tyms buds.

Target Expression Profiles Based on Transcriptome Sequencing

To profile the expression of the targets, a transcriptome sequencing project was performed separately on buds of 231–330 and Tyms. A total of 25,015,186 and 24,295,576 clean reads were generated from the 231–330 and Tyms buds, respectively. Approximately 76.4 and 72.6% of the reads were mapped to the reference genome and 59.3 and 57.9% of the reads were mapped to the reference gene sets of *B. rapa*, respectively. The expression of the target genes was called from the transcriptome data for further analysis. A total of 367 and 352 out of the 376 target genes were expressed in the 231–330 and Tyms buds, respectively. Among them, 44 were up-regulated and 32 were down-regulated (two-fold change and FDR $< 10^{-3}$) in Tyms compared with 231–330. Most of the differentially expressed targets were pollen development-related genes according to the annotation (**Tables 4, 5**).

TABLE 4 | The up-regulated targets in Tyms compared to 231-330 buds.

geneID	RPKM (231–330)	RPKM (Tyms)	log2 (T/231)	FDR	Annotation
Bra011781	0.153	2.406	3.980	1.13E−10	Growth-regulating factor 2J
Bra004474	0.881	8.341	3.243	1.42E−22	Galactinol synthase
Bra006956	0.420	1.767	2.073	4.46E−04	Growth-regulating factor 4
Bra017379	3.814	14.809	1.957	6.43E−16	Lysosomal thiol reductase family protein
Bra012155	3.935	14.720	1.904	2.77E−41	U-box domain-containing protein 2
Bra004363	1.826	6.322	1.791	3.02E−08	Squamosa promoter-binding protein-like 6
Bra004892	62.505	215.114	1.783	0	Transducin family protein
Bra000864	0.256	0.860	1.751	8.10E−04	Ent-copalyl diphosphate synthase
Bra027897	20.504	67.391	1.717	8.67E−123	CYP96A15
Bra028175	42.722	135.558	1.666	3.16E−147	Ribulose bisphosphate carboxylase small chain
Bra017851	0.872	2.751	1.658	3.44E−05	Growth-regulating factor 2P
Bra005019	30.610	96.012	1.649	8.96E−157	RCA (RUBISCO activase)
Bra000096	2.553	7.910	1.631	4.77E−14	Calmodulin-binding protein-like
Bra022685	1.998	5.976	1.580	5.35E−08	Cup-shaped cotyledon2
Bra029216	2.231	6.648	1.575	1.67E−07	LRR-repeat protein
Bra016937	14.438	41.059	1.508	6.99E−34	S-adenosyl-L-methionine-dependent methyltransferase
Bra017143	2.326	6.533	1.490	2.49E−15	Ceramidase family protein
Bra012263	44.469	120.598	1.439	3.82E−68	Zinc-binding family protein
Bra001369	3.402	9.054	1.412	4.47E−09	3-phosphoinositide-dependent protein kinase-1
Bra031329	37.673	98.231	1.383	8.45E−89	Cell wall-plasma membrane linker protein
Bra028181	2288.119	5965.931	1.383	0	Ribulose bisphosphate carboxylase small chain
Bra023844	5.298	13.778	1.379	5.22E−24	CRK10 (cysteine-rich rlk10)
Bra023235	89.500	228.685	1.353	2.52E−232	ATP binding / phosphoribulokinase
Bra034079	2.458	6.219	1.339	2.56E−16	Disease resistance protein
Bra030820	4.925	12.084	1.295	3.46E−10	NAC1
Bra024616	9.169	22.415	1.290	2.68E−21	Oxidoreductase
Bra030727	23.394	56.454	1.271	1.36E−51	Calreticulin-3
Bra013767	2.533	5.912	1.223	9.3E−06	Growth-regulating factor 8
Bra005131	47.935	111.401	1.217	1.35E−112	cax51
Bra030188	6.137	14.006	1.190	4.11E−06	Nudix hydrolase 25
Bra025427	1.529	3.475	1.185	2.49E−04	Brassinosteroid-6-oxidase 1
Bra036417	4.213	9.514	1.175	2.08E−17	Disease resistance protein
Bra013013	15.281	33.823	1.146	1.46E−26	Calcium sensing receptor
Bra012691	2.217	4.896	1.143	5.87E−04	2OG-Fe(II) oxygenase family protein
Bra020262	3.260	7.194	1.142	5.32E−07	TOE2 transcription factor
Bra023066	1.993	4.257	1.095	7.54-04	Growth-regulating factor 4
Bra014184	7.485	15.606	1.060	4.65E−15	Cytochrome c biogenesis protein CCS1
Bra003311	4.714	9.724	1.045	2.85E−09	Scarecrow transcription factor
Bra040873	42.657	87.512	1.037	1.49E−64	Aminoacylase
Bra012008	3.631	7.428	1.033	1.17E−06	Cysteine/histidine-rich C1 domain-containing
Bra036135	8.977	18.315	1.029	7.59E−05	Lysosome-related
Bra011537	56.705	114.452	1.013	9.85E−94	O-Glycosyl hydrolases family 17 protein
Bra013556	3.991	8.044	1.011	7.72E−10	Homeobox-leucine zipper protein HDG2-like
Bra003518	4.133	8.286	1.004	9.55E−09	Transport inhibitor response 1

miRNA-target Network Underlying Ogura-CMS

The differentially expressed target genes were targeted by 10 known miRNAs (miR159, miR164, miR171, miR172, miR393, miR396, miR397, miR1885, miR5654, and miR6034) and 33 novel miRNAs. However, none of these known miRNAs were significantly differentially expressed according to the two-fold criteria. Based on the combined analysis of the expression profiles of miRNAs and their targets without the significant test filters, the expression patterns of the miRNA-targets could be classified into four clusters. As shown in **Figure 5A**, cluster I contained more than half of the miRNA-target pairs, in which

TABLE 5 | The down-regulated targets in Tyms compared to 231-330 buds.

geneID	RPKM (231-330)	RPKM (Tyms)	log2 (T/231)	FDR	Annotation
Bra016846	19.369	0.001	−14.241	1.71E−72	SYP125
Bra031759	18.677	0.001	−14.189	6.8E−70	SYP125
Bra019872	12.375	0.162	−6.258	9.21E−43	SYP125
Bra015756	9.216	0.001	−13.170	3.18E−44	Serine/threonine-protein kinase
Bra008330	5.324	0.001	−12.378	2.81E−17	Pleckstrin homology (PH) domain-containing
Bra026884	2.748	0.001	−11.424	7.39E−16	Pentatricopeptide (PPR) repeat-containing
Bra035163	547.451	1.757	−8.284	0	H^+-ATPase 9
Bra013168	677.150	6.843	−6.629	0	H^+-ATPase 6
Bra003491	331.459	1.268	−8.030	0	VGDH2 (Vanguard 1 homolog 2)
Bra004481	1608.187	8.885	−7.500	0	Pectinesterase 5
Bra030518	19.054	0.164	−6.864	3.32E−133	Clathrin assembly protein
Bra030371	607.258	10.827	−5.810	0	Late embryogenesis abundant protein−like
Bra005178	226.835	4.629	−5.615	0	Actin 3
Bra028534	28.543	0.607	−5.555	6.96E−138	Purple acid phosphatase 21
Bra039006	2.246	0.052	−5.419	7.14E−23	Retroelement pol polyprotein
Bra022826	12.903	0.408	−4.984	2.66E−56	Triacylglycerol lipase
Bra033779	3.376	0.115	−4.882	6.72E−08	HPP integral membrane domain−containing
Bra029964	55.449	2.248	−4.624	1.35E−191	2−Oxoglutarate-Fe(II)-dependent oxygenase like
Bra005140	118.770	5.278	−4.492	0	Laccase-4
Bra001912	3.939	0.196	−4.331	1.54E−28	Alcohol oxidase-related
Bra026378	3.795	0.222	−4.097	1.68E−23	COBra-like protein 11 precursor
Bra007991	1150.764	85.857	−3.745	0	Sucrose transporter SUC1
Bra031453	1.928	0.219	−3.137	4.76E−07	Serine carboxypeptidase S10 family protein
Bra020052	14.631	2.603	−2.491	2.66E−53	Plasma membrane sulfate transporter
Bra015685	7.629	1.655	−2.205	3.7E−19	Sugar transporter
Bra036819	11.142	3.174	−1.812	2.07E−09	Putative ROP family GTPase
Bra033251	24.180	7.172	−1.753	2.46E−50	CYP86A4
Bra032722	13.487	4.457	−1.598	2.09E−25	Phosphoinositide phospholipase
Bra039028	75.548	28.242	−1.420	8.75E−108	Nucleotide-diphospho-sugar transferase like
Bra013701	3.431	1.362	−1.333	5.71E−04	F−box family protein
Bra000531	16.656	7.927	−1.071	1.35E−13	MYB81
Bra012414	24.322	11.883	−1.033	8.74E−06	Bet v I allergen family protein

miRNAs were down-regulated and the targets were up-regulated in Tyms. Approximately 16 and 24% of the miRNA-target pairs were synergistically down-regulated (cluster II) or up-regulated (cluster III), respectively. Less than 8% of the miRNAs were up-regulated while their targets were down-regulated. These results indicate the miRNAs provide an efficient buffer system for fine-tuning the expression of genes involved in bud and pollen development. When a cut-off of >1.5-fold change was used, seven miRNAs were down-regulated and released the expression of 18 targets in Tyms. In contrast, four miRNAs were up-regulated and suppressed the expression of five targets (**Figure 5B**). miR156 and *SPLs* represent a cascade of controls for many important developmental processes, including the proper development of sporogenic tissues of the anther (Xing et al., 2010). miR156 and *SPLs* are expressed in high abundance in many tissues, including buds (Zhang et al., 2011b). Our present study found that miR156 and *SPLs* were down- or up-regulated in Tyms buds by more than 1.5-fold but less than two-fold. It is possible that the expression of these genes was changed more significantly in sporogenic tissues but was submerged by the expression in other tissues of the buds. Using a two-fold threshold, three and two novel miRNAs were up-regulated in 231–330 and Tyms, resulting in the down-regulation of five and three of their target genes in the corresponding tissues, respectively (**Figure 5C**). Among them, a cytochrome c biogenesis protein was targeted by novel-miR-180, a Ca^{2+}/H^+ antiporter was targeted by novel-miR-191, a P450 and a glycogenin glucosyltransferase (GGT) were targeted by novel-miR-6; all of these genes have been implicated in pollen development by previous studies (Balk and Leaver, 2001; Welchen and Gonzalez, 2005; Morant et al., 2007; Song et al., 2009; Li et al., 2010, 2012; Rennie et al., 2014) and were up-regulated in Tyms in our analysis. Interestingly, novel-miR-448 and novel-miR-335 were specifically expressed in Tyms; this is especially worth noticing because the novel-miR-335 was expressed at a relatively high level. The expression of its

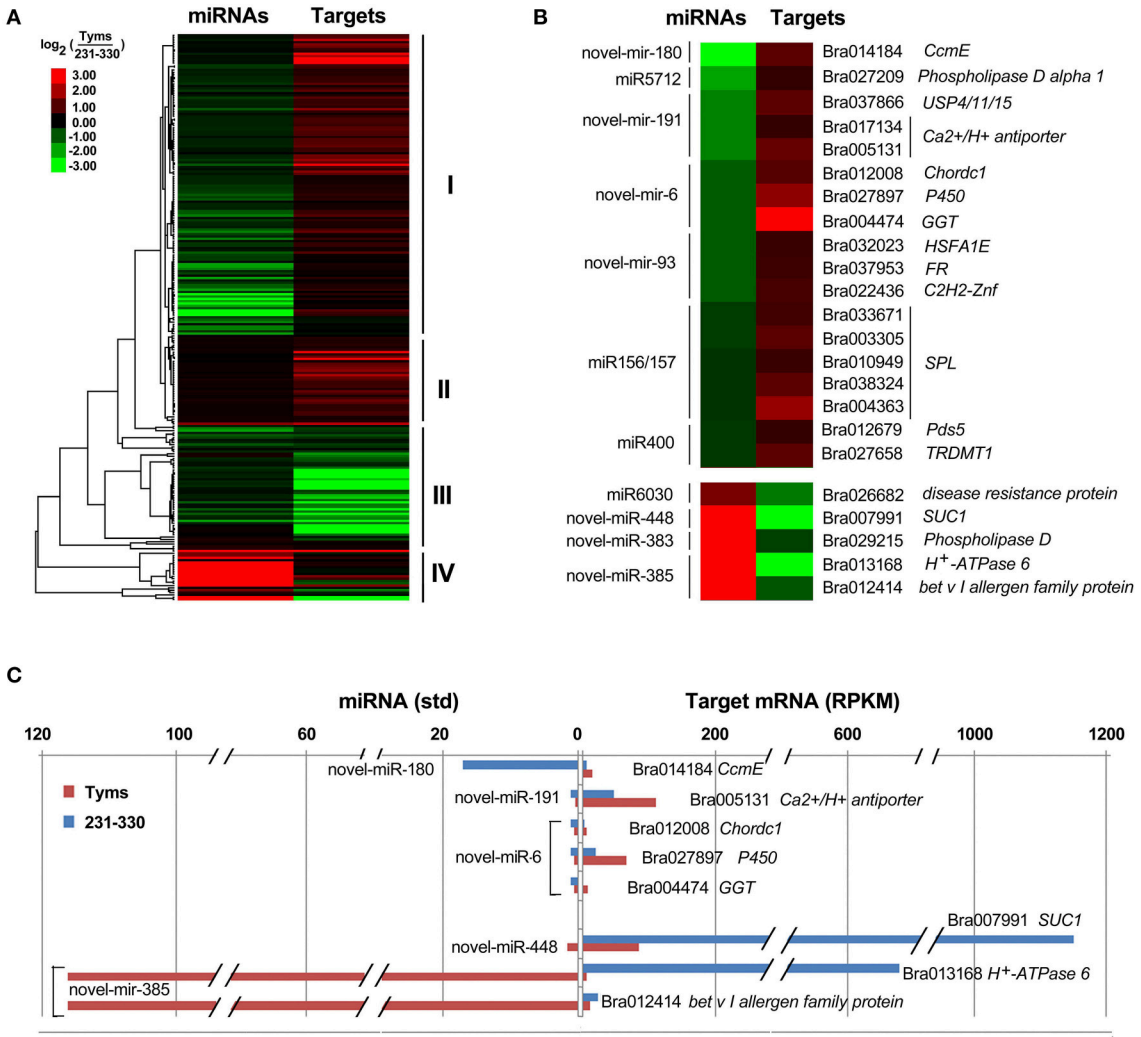


FIGURE 5 | Expression profiles of miRNA-targets. (A) Heatmap of overall expression patterns of miRNA-targets. **(B)** Heatmap of expression levels of miRNAs and their targets with more than 1.5-fold changes between Tyms and 231–330. **(C)** Bar chart of expression levels of miRNAs and their targets with more than two-fold changes between Tyms and 231–330. CcmE, cytochrome c biogenesis protein; USP4/11/15, ubiquitin carboxyl-terminal hydrolase 4/11/15; Chordc 1, cysteine/histidine-rich C1 domain-containing protein; P450, cytochrome P450; GGT, glycogenin glucosyltransferase; FR, ferric reduction oxidase 7; C2H2-Znf, C2H2 zinc finger protein 1; SPL, SQUAMOSA promoter binding protein-like; Pds5, sister chromatid cohesion protein PDS5; TRDMT1, tRNA (cytosine38-C5)-methyltransferase; SUC1, sucrose transporter.

target genes (*sucrose transporter SUC1* and *H^+-ATPase 6*) was high in 231–330 but suppressed by approximately 100-fold in Tyms. SUC1 is an enzyme that catalyzes the degradation of sucrose and performs important roles for pollen germination in *Arabidopsis* (Sivitz et al., 2008). Thus, the down-regulation of *SUC1* may result in energy deficiency in Tyms buds and thus abort pollen development. A study in *Nicotiana plumbaginifolia* reported that co-suppression of an H^+-ATPase impaired sucrose translocation and male fertility (Zhao et al., 2000). In our present study, both a *SUC1* and an *H^+-ATPase* were indicated as targets of two novel miRNAs and were sufficiently suppressed by these two miRNAs in the Ogura-CMS buds (**Figure 5C**). This finding indicated that the novel-miR-335/*H^+-ATPase* and novel-miR-448/*SUC1* cascade could play important roles in male sterility in Ogura-CMS. The secondary structure of the novel-miR-335 and novel-miR-448 and the degradation map of *H^+-ATPase* and *SUC1* were shown in **Figure 6**. The suppression of *H^+-ATPase* and *SUC1* expression were validated by Q-PCR analysis (**Figure 7**). This finding filled a gap within the cross-talk network between the orf138 locus in mitochondria and the effecter genes in the chromosome. However, more studies are still needed to linkup the orf138 locus and miRNA regulatory networks.

CONCLUSION

The present study used deep sequencing technology and performed a series of whole genome small RNA, degradome,

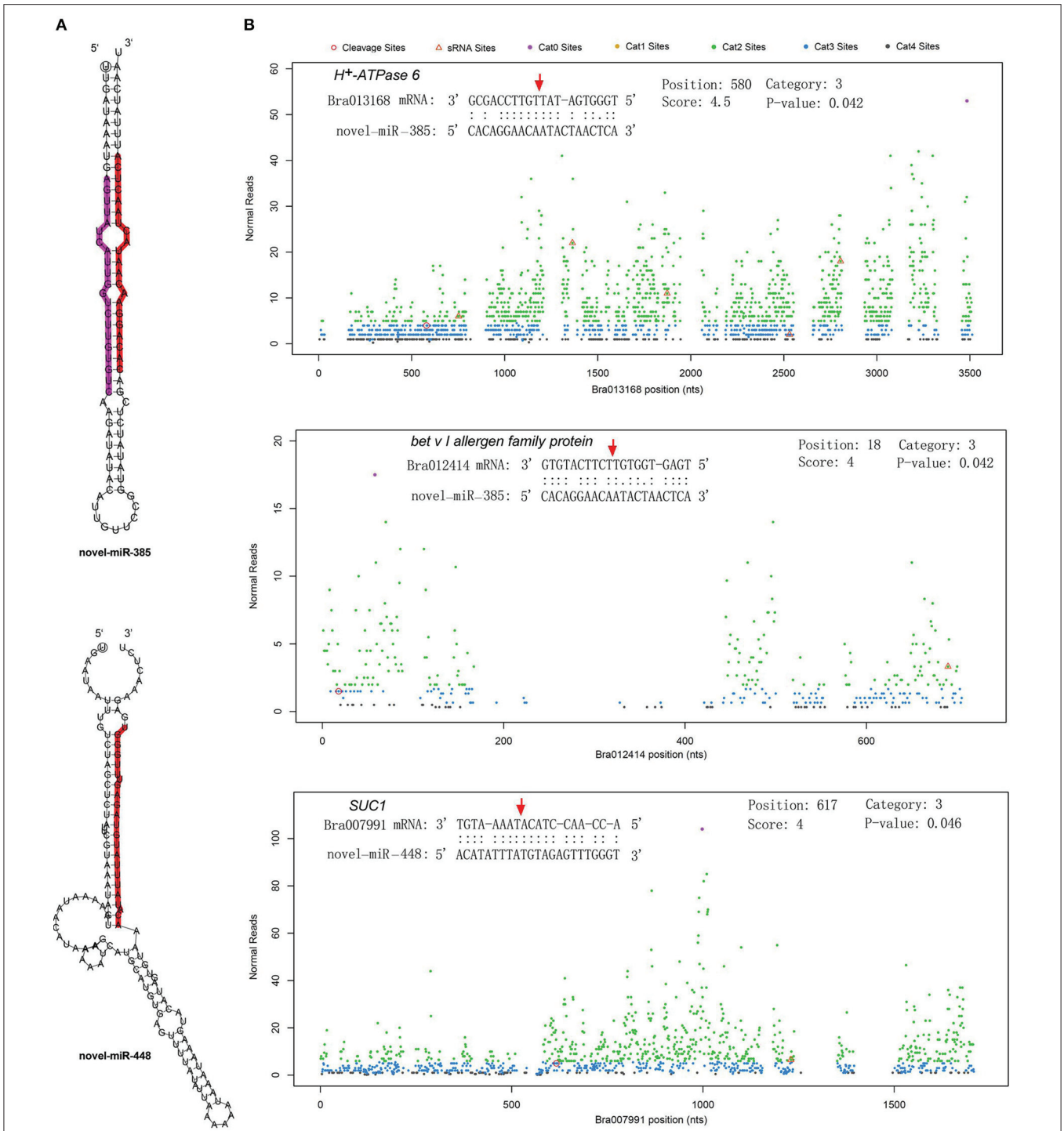


FIGURE 6 | Secondary structure of two novel miRNAs and the cleavage plots of their targets. (A) Secondary structure of pre-miRNA hairpins. Red shaded areas indicate the dominant mature miRNAs, violet shaded areas indicates the reverse complementary mature miRNAs. **(B)** Cutting plots of miRNA targets confirmed using degradome sequencing. The corresponding miRNA:mRNA alignments are shown on the top. The red arrows indicate the miRNA-directed cleavage positions. The y-axis shows the nucleotide position in the target gene. The x-axis indicates the number of cleaved ends detected in the degradome analysis.

and transcriptome analyses on Chinese cabbage buds from both Ogura-CMS and its maintainer. A total of 289 (69 families) known and 426 novel miRNAs were identified, which was much higher than the number of miRNAs found in the buds of the closely relative subspecies *B. campestris* ssp. *chinensis* (Jiang et al., 2014). These miRNAs not only validated the finding in the

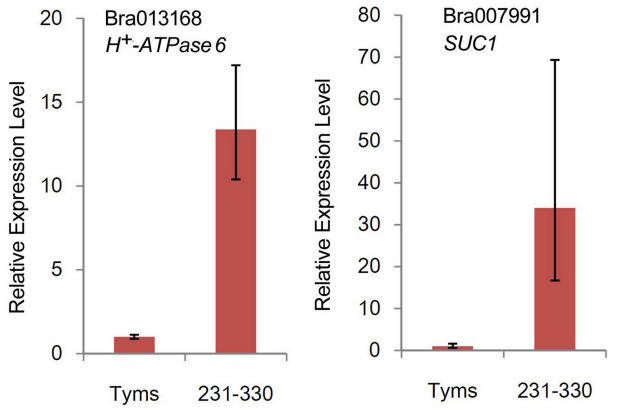


FIGURE 7 | Relative expression levels of two target genes revealed by Q-PCR.

previous study on *B. campestris* ssp. *chinensis* but also contained many novel miRNAs that were not previously reported. A large number of targets were firstly validated in this study. The combinational profiling of miRNAs and targets revealed a regulatory network contributing to bud development, especially for pollen engenderation. The finding of two novel miRNA/target cascades (novel-miR-335/H^+-*ATPase* and novel-miR-448/*SUC1*) will provide new insights into the communication between the mitochondria and chromosome and take one step toward filling in the gap in the regulatory network mechanism from the orf138 locus to the pollen abortion in the Ogura-CMS plants. The true functions of these two miRNAs warrant more solid experiments, such as the transgenic complementary test.

AUTHOR CONTRIBUTIONS

XW and XWZ designed the study. XW performed the experiments. XW and XHZ analyzed the data and drafted the manuscript. YY and QZ assisted with the bioinformatics analysis and aided in writing the manuscript. QY, ZW, YZ, and WJ aided in performing the experiments. XL and FW modified the manuscript. All of the authors carefully checked and approved this version of the manuscript.

DATA ACCESS

RNAseq are submitted to EMBL/NCBI/SRA with the accession numbers SRR2132359, SRR2136647, SRR2149955, SRR2132463, SRR2136646, SRR2149956.

FUNDING

This work was supported by funding from North Henan Station, the National Vegetable Industry Technology System (CARS-25-27), Excellent technology innovation team of Henan province, the National High Technology Research and Development Program of China (863 Program) (2012AA100202-7), The National Key Technology R&D Program of China (2012BAD02B01-3).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00894

REFERENCES

Addo-Quaye, C., Miller, W., and Axtell, M. J. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics* 25, 130–131. doi: 10.1093/bioinformatics/btn604

Aukerman, M. J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *Plant Cell* 15, 2730–2741. doi: 10.1105/tpc.016238

Balk, J., and Leaver, C. J. (2001). The PET1-CMS mitochondrial mutation in sunflower is associated with premature programmed cell death and cytochrome c release. *Plant Cell* 13, 1803–1818. doi: 10.1105/tpc.13.8.1803

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi: 10.1016/S0092-8674(04)00045-5

Baulcombe, D. (2004). RNA silencing in plants. *Nature* 431, 356–363. doi: 10.1038/nature02874

Bhogale, S., Mahajan, A. S., Natarajan, B., Rajabhoj, M., Thulasiram, H. V., and Banerjee, A. K. (2014). miRNA156-A potential graft-transmissible microRNA that modulates plant architecture and tuberization in potato (*Solanum tuberosum* ssp *andigena*). *Plant Physiol.* 164, 1011–1027. doi: 10.1104/pp.113.230714

Carrington, J. C., and Ambros, V. (2003). Role of microRNAs in plant and animal development. *Science* 301, 336–338. doi: 10.1126/science.1085242

Chambers, C., and Shuai, B. (2009). Profiling microRNA expression in Arabidopsis pollen using microRNA array and real-time PCR. *BMC Plant Biol.* 9:87. doi: 10.1186/1471-2229-9-87

Chase, C. D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial–nuclear interactions. *Trends Genet.* 23, 81–90. doi: 10.1016/j.tig.2006.12.004

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* 303, 2022–2025. doi: 10.1126/science.1088060

Desloire, S., Gherbi, H., Laloui, W., Marhadour, S., Clouet, V., Cattolico, L., et al. (2003). Identification of the fertility restoration locus, *Rfo*, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep.* 4, 588–594. doi: 10.1038/sj.embor.embor848

Dettmer, J., Schubert, D., Calvo-Weimar, O., Stierhof, Y. D., Schmidt, R., and Schumacher, K. (2005). Essential role of the V-ATPase in male gametophyte development. *Plant J.* 41, 117–124. doi: 10.1111/j.1365-313X.2004.02282.x

Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* 38, W64–W70. doi: 10.1093/nar/gkq310

Floyd, S. K., and Bowman, J. L. (2004). Gene regulation: ancient microRNA target sequences in plants. *Nature* 428, 485–486. doi: 10.1038/428485a

Folkes, L., Moxon, S., Woolfenden, H. C., Stocks, M. B., Szittya, G., Dalmay, T., et al. (2012). PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. *Nucleic Acids Res.* 40:e103. doi: 10.1093/nar/gks277

Fujii, S., and Small, I. (2011). The evolution of RNA editing and pentatricopeptide repeat genes. *New phytol.* 191, 37–47. doi: 10.1111/j.1469-8137.2011.03746.x

German, M. A., Luo, S., Schroth, G., Meyers, B. C., and Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* 4, 356–362. doi: 10.1038/nprot.2009.8

Grant-Downton, R., Hafidh, S., Twell, D., and Dickinson, H. G. (2009a). Small RNA pathways are present and functional in the angiosperm male gametophyte. *Mol. Plant* 2, 500–512. doi: 10.1093/mp/ssp003

Grant-Downton, R., Le Trionnaire, G., Schmid, R., Rodriguez-Enriquez, J., Hafidh, S., Mehdi, S., et al. (2009b). MicroRNA and tasiRNA diversity in mature pollen of *Arabidopsis thaliana*. *BMC Genomics* 10:643. doi: 10.1186/1471-2164-10-643

Grelon, M., Budar, F., Bonhomme, S., and Pelletier, G. (1994). Ogura cytoplasmic male-sterility (CMS)-associated orf138 is translated into a mitochondrial membrane polypeptide in male-sterile Brassica cybrids. *Mol. Gen. Genet.* 243, 540–547. doi: 10.1007/BF00284202

Guo, H. S., Xie, Q., Fei, J. F., and Chua, N. H. (2005). MicroRNA directs mRNA cleavage of the transcription factor *NAC1* to downregulate auxin signals for arabidopsis lateral root development. *Plant Cell* 17, 1376–1386. doi: 10.1105/tpc.105.030841

He, H., Liang, G., Li, Y., Wang, F., and Yu, D. (2014). Two young miRNAs originating from target duplication mediate nitrogen starvation adaptation via regulation of glucosinolate synthesis in *Arabidopsis thaliana*. *Plant Physiol.* 164, 853–865. doi: 10.1104/pp.113.228635

Jiang, J., Lv, M., Liang, Y., Ma, Z., and Cao, J. (2014). Identification of novel and conserved miRNAs involved in pollen development in *Brassica campestris* ssp. chinensis by high-throughput sequencing and degradome analysis. *BMC Genomics* 15:146. doi: 10.1186/1471-2164-15-146

Juarez, M. T., Kui, J. S., Thomas, J., Heller, B. A., and Timmermans, M. C. (2004). microRNA-mediated repression of *rolled leaf1* specifies maize leaf polarity. *Nature* 428, 84–88. doi: 10.1038/nature02363

Katiyar-Agarwal, S., and Jin, H. (2010). Role of small RNAs in host-microbe interactions. *Annu. Rev. Phytopathol.* 48, 225. doi: 10.1146/annurev-phyto-073009-114457

Khraiwesh, B., Arif, M. A., Seumel, G. I., Ossowski, S., Weigel, D., Reski, R., et al. (2010). Transcriptional control of gene expression by microRNAs. *Cell* 140, 111–122. doi: 10.1016/j.cell.2009.12.023

Kim, B., Yu, H. J., Park, S. G., Shin, J. Y., Oh, M., Kim, N., et al. (2012). Identification and profiling of novel microRNAs in the *Brassica rapa* genome based on small RNA deep sequencing. *BMC Plant Biol.* 12:218. doi: 10.1186/1471-2229-12-218

Kinoshita, N., Wang, H., Kasahara, H., Liu, J., MacPherson, C., Machida, Y., et al. (2012). IAA-Ala Resistant3, an evolutionarily conserved target of miR167, mediates *Arabidopsis* root architecture changes during high osmotic stress. *Plant Cell* 24, 3590–3602. doi: 10.1105/tpc.112.097006

Krishnasamy, S., and Makaroff, C. A. (1993). Characterization of the radish mitochondrial orfB locus: possible relationship with male sterility in Ogura radish. *Curr. Genet.* 24, 156–163. doi: 10.1007/BF00324680

Krishnasamy, S., and Makaroff, C. A. (1994). Organ-specific reduction in the abundance of a mitochondrial protein accompanies fertility restoration in cytoplasmic male-sterile radish. *Plant Mol. Biol.* 26, 935–946. doi: 10.1007/BF00028860

Kurihara, Y., and Watanabe, Y. (2004). *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12753–12758. doi: 10.1073/pnas.0403115101

Li, H. M., Chen, H., Yang, Z. N., and Gong, J. M. (2012). Cdi gene is required for pollen germination and tube growth in *Arabidopsis*. *FEBS Lett.* 586, 1027–1031. doi: 10.1016/j.febslet.2012.02.046

Li, H., Pinot, F., Sauveplane, V., Werck-Reichhart, D., Diehl, P., Schreiber, L., et al. (2010). Cytochrome P450 family member CYP704B2 catalyzes the ω-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. *Plant Cell* 22, 173–190. doi: 10.1105/tpc.109.070326

Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: short oligonucleotide alignment program. *Bioinformatics* 24, 713–714. doi: 10.1093/bioinformatics/btn025

Liu, G., Tian, H., Huang, Y. Q., Hu, J., Ji, Y. X., Li, S. Q., et al. (2012). Alterations of mitochondrial protein assembly and jasmonic acid biosynthesis pathway in Honglian (HL)-type cytoplasmic male sterility rice. *J. Biol. Chem.* 287, 40051–40060. doi: 10.1074/jbc.M112.382549

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Luo, D., Xu, H., Liu, Z., Guo, J., Li, H., Chen, L., et al. (2013). A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nat. Genet.* 45, 573–577. doi: 10.1038/ng.2570

Morant, M., Jørgensen, K., Schaller, H., Pinot, F., Møller, B. L., Werck-Reichhart, D., et al. (2007). CYP703 is an ancient cytochrome P450 in land plants catalyzing in-chain hydroxylation of lauric acid to provide building blocks for sporopollenin synthesis in pollen. *Plant Cell* 19, 1473–1487. doi: 10.1105/tpc.106.045948

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Nagpal, P., Ellis, C. M., Weber, H., Ploense, S. E., Barkawi, L. S., Guilfoyle, T. J., et al. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132, 4107–4118. doi: 10.1242/dev.01955

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., et al. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* 312, 436–439. doi: 10.1126/science.1126088

Ogura, H. (1968). Studies on the new male-sterility in Japanese radish, with special reference to the utilization of this sterility towards the practical raising of hybrid seeds. *Mem. Fac. Agric. Kagoshima Univ.*6, 39–78.

Pelletier, G., Primard, C., Vedel, F., Chetrit, P., Remy, R., and Renard, M. (1983). Intergeneric cytoplasmic hybridization in Cruciferae by protoplast fusion. *Mol. Gen. Genet.* 191, 244–250. doi: 10.1007/BF00334821

Rennie, E. A., Ebert, B., Miles, G. P., Cahoon, R. E., Christiansen, K. M., Stonebloom, S., et al. (2014). Identification of a sphingolipid alpha-glucuronosyltransferase that is essential for pollen function in *Arabidopsis*. *Plant Cell* 26, 3314–3325. doi: 10.1105/tpc.114.129171

Romualdi, C., Bortoluzzi, S., D'Alessi, F., and Danieli, G. A. (2003). IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genomics* 12, 159–162. doi: 10.1152/physiolgenomics.00096.2002

Si-Ammour, A., Windels, D., Arn-Bouldoires, E., Kutter, C., Ailhas, J., Meins, F. Jr., et al. (2011). miR393 and secondary siRNAs regulate expression of the TIR1/AFB2 auxin receptor clade and auxin-related development of *Arabidopsis* leaves. *Plant Physiol.* 157, 683–691. doi: 10.1104/pp.111.180083

Sivitz, A. B., Reinders, A., and Ward, J. M. (2008). *Arabidopsis* sucrose transporter AtSUC1 is important for pollen germination and sucrose-induced anthocyanin accumulation. *Plant Physiol.* 147, 92–100. doi: 10.1104/pp.108.118992

Song, L. F., Zou, J. J., Zhang, W. Z., Wu, W. H., and Wang, Y. (2009). Ion transporters involved in pollen germination and pollen tube tip-growth. *Plant Signal. Behav.* 4, 1193–1195. doi: 10.4161/psb.4.12.10063

Sunkar, R., Chinnusamy, V., Zhu, J., and Zhu, J.-K. (2007). Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. *Trends Plant Sci.* 12, 301–309. doi: 10.1016/j.tplants.2007.05.001

Tarutani, Y., Shiba, H., Iwano, M., Kakizaki, T., Suzuki, G., Watanabe, M., et al. (2010). Trans-acting small RNA determines dominance relationships in *Brassica* self-incompatibility. *Nature* 466, 983–986. doi: 10.1038/nature09308

Terrile, M. C., París, R., Calderón-Villalobos, L. I., Iglesias, M. J., Lamattina, L., Estelle, M., et al. (2012). Nitric oxide influences auxin signaling through S-nitrosylation of the *Arabidopsis* TRANSPORT INHIBITOR RESPONSE 1 auxin receptor. *Plant J.* 70, 492–500. doi: 10.1111/j.1365-313X.2011.04885.x

Vaucheret, H., Vazquez, F., Crete, P., and Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* 18, 1187–1197. doi: 10.1101/gad.1201404

Wang, J. W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell* 138, 738–749. doi: 10.1016/j.cell.2009.06.014

Wang, J. W., Park, M. Y., Wang, L. J., Koo, Y., Chen, X. Y., Weigel, D., et al. (2011a). miRNA control of vegetative phase change in trees. *PLoS Genet.* 7:e1002012. doi: 10.1371/journal.pgen.1002012

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011b). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Wang, Z. W., Zhang, Y. J., Xiang, C. P., Mei, S. Y., Zhou, Y., Chen, G. P., et al. (2008). A new fertility restorer locus linked closely to the Rfo locus for cytoplasmic male sterility in radish. *Theor. Appl. Genet.* 117, 313–320. doi: 10.1007/s00122-008-0776-5

Wei, L. Q., Yan, L. F., and Wang, T. (2011). Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. *Genome Biol.* 12:R53. doi: 10.1186/gb-2011-12-6-r53

Welchen, E., and Gonzalez, D. H. (2005). Differential expression of the *Arabidopsis* cytochrome c genes *Cytc-1* and *Cytc-2*. Evidence for the involvement of TCP-domain protein-binding elements in anther- and meristem-specific expression of the *Cytc-1* gene. *Plant Physiol.* 139, 88–100. doi: 10.1104/pp.105.065920

Widemann, E., Miesch, L., Lugan, R., Holder, E., Heinrich, C., Aubert, Y., et al. (2013). The amidohydrolases IAR3 and ILL6 contribute to jasmonoyl-isoleucine hormone turnover and generate 12-hydroxyjasmonic acid upon wounding in *Arabidopsis* leaves. *J. Biol. Chem.* 288, 31701–31714. doi: 10.1074/jbc.M113.499228

Woo, M. O., Ham, T. H., Ji, H. S., Choi, M. S., Jiang, W., Chu, S. H., et al. (2008). Inactivation of the *UGPase1* gene causes genic male sterility and endosperm chalkiness in rice (*Oryza sativa* L.). *Plant J.* 54, 190–204. doi: 10.1111/j.1365-313X.2008.03405.x

Woodson, J. D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. *Nat. Rev. Genet.* 9, 383–395. doi: 10.1038/nrg2348

Wu, G., Park, M. Y., Conway, S. R., Wang, J. W., Weigel, D., and Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138, 750–759. doi: 10.1016/j.cell.2009.06.031

Xian, Z., Huang, W., Yang, Y., Tang, N., Zhang, C., Ren, M., et al. (2014). miR168 influences phase transition, leaf epinasty, and fruit development via *SlAGO1s* in tomato. *J. Exp. Bot.* 65, 6655–6666. doi: 10.1093/jxb/eru387

Xing, S., Salinas, M., Garcia-Molina, A., Hohmann, S., Berndtgen, R., and Huijser, P. (2013). *SPL8* and miR156-targeted *SPL* genes redundantly regulate Arabidopsis gynoecium differential patterning. *Plant J.* 75, 566–577. doi: 10.1111/tpj.12221

Xing, S., Salinas, M., Hohmann, S., Berndtgen, R., and Huijser, P. (2010). miR156-targeted and nontargeted *SBP-box* transcription factors act in concert to secure male fertility in Arabidopsis. *Plant Cell* 22, 3935–3950. doi: 10.1105/tpc.110.079343

Yamagishi, H., and Bhat, S. R. (2014). Cytoplasmic male sterility in Brassicaceae crops. *Breed. Sci.* 64, 38–47. doi: 10.1270/jsbbs.64.38

Yamasaki, H., Abdel-Ghany, S. E., Cohu, C. M., Kobayashi, Y., Shikanai, T., and Pilon, M. (2007). Regulation of copper homeostasis by micro-RNA in *Arabidopsis*. *J. Biol. Chem.* 282, 16369–16378. doi: 10.1074/jbc.M700138200

Yasumoto, K., Terachi, T., and Yamagishi, H. (2009). A novel Rf gene controlling fertility restoration of Ogura male sterility by RNA processing of orf138 found in Japanese wild radish and its STS markers. *Genome* 52, 495–504. doi: 10.1139/G09-026

Yu, S., Galvão, V. C., Zhang, Y. C., Horrer, D., Zhang, T. Q., Hao, Y. H., et al. (2012). Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. *Plant Cell* 24, 3320–3332. doi: 10.1105/tpc.112.101014

Zhang, X., Zou, Z., Gong, P., Zhang, J., Ziaf, K., Li, H., et al. (2011a). Over-expression of microRNA169 confers enhanced drought tolerance to tomato. *Biotechnol. Lett.* 33, 403–409. doi: 10.1007/s10529-010-0436-0

Zhang, X., Zou, Z., Zhang, J., Zhang, Y., Han, Q., Hu, T., et al. (2011b). Over-expression of sly-miR156a in tomato results in multiple vegetative and reproductive trait alterations and partial phenocopy of the sft mutant. *FEBS Lett.* 585, 435–439. doi: 10.1016/j.febslet.2010.12.036

Zhao, R., Dielen, V., Kinet, J. M., and Boutry, M. (2000). Cosuppression of a plasma membrane H(+)-ATPase isoform impairs sucrose translocation, stomatal opening, plant growth, and male fertility. *Plant Cell* 12, 535–546. doi: 10.1105/tpc.12.4.535

Zhou, C. M., Zhang, T. Q., Wang, X., Yu, S., Lian, H., Tang, H., et al. (2013). Molecular basis of age-dependent vernalization in *Cardamine flexuosa*. *Science* 340, 1097–1100. doi: 10.1126/science.1234340

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Wei, Zhang, Yao, Yuan, Li, Wei, Zhao, Zhang, Wang, Jiang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 26 July 2016
doi: 10.3389/fpls.2016.01049

Cytogenetic Diversity of Simple Sequences Repeats in Morphotypes of *Brassica rapa* ssp. *chinensis*

Jin-shuang Zheng[1,2], Cheng-zhen Sun[1,2], Shu-ning Zhang[1], Xi-lin Hou[1] and Guusje Bonnema[3]*

[1] State Key Laboratory of Crop Genetics and Germplasm Enhancement, Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in East China, Ministry of Horticulture, Nanjing Agricultural University, Nanjing, China, [2] Hebei Normal University of Science and Technology, Qinhuangdao, China, [3] Wageningen UR Plant Breeding, Wageningen University and Research Centre, Wageningen, Netherlands

A significant fraction of the nuclear DNA of all eukaryotes is comprised of simple sequence repeats (SSRs). Although these sequences are widely used for studying genetic variation, linkage mapping and evolution, little attention had been paid to the chromosomal distribution and cytogenetic diversity of these sequences. In this paper, we report the distribution characterization of mono-, di-, and tri-nucleotide SSRs in *Brassica rapa* ssp. *chinensis*. Fluorescence *in situ* hybridization was used to characterize the cytogenetic diversity of SSRs among morphotypes of *B. rapa* ssp. *chinensis*. The proportion of different SSR motifs varied among morphotypes of *B. rapa* ssp. *chinensis*, with tri-nucleotide SSRs being more prevalent in the genome of *B. rapa* ssp. *chinensis*. We determined the chromosomal locations of mono-, di-, and tri-nucleotide repeat loci. The results showed that the chromosomal distribution of SSRs in the different morphotypes is non-random and motif-dependent, and allowed us to characterize the relative variability in terms of SSR numbers and similar chromosomal distributions in centromeric/peri-centromeric heterochromatin. The differences between SSR repeats with respect to abundance and distribution indicate that SSRs are a driving force in the genomic evolution of *B. rapa* species. Our results provide a comprehensive view of the SSR sequence distribution and evolution for comparison among morphotypes *B. rapa* ssp. *chinensis*.

Keywords: *Brassica rapa*, simple sequence repeats, fluorescence *in situ* hybridization, cytogenetic diversity, heterochromatin

OPEN ACCESS

Edited by:
Sarvajeet Singh Gill, Maharshi Dayanand University, India

Reviewed by:
Maoteng Li, Huazhong University of Science and Technology, China
Santosh Tiwari, Maharshi Dayanand University, India

****Correspondence:***
Xi-lin Hou
hxl@njau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *27 February 2016*
Accepted: *04 July 2016*
Published: *26 July 2016*

Citation:
Zheng J-s, Sun C-z, Zhang S-n, Hou X-l and Bonnema G (2016) Cytogenetic Diversity of Simple Sequences Repeats in Morphotypes of Brassica rapa ssp. chinensis. Front. Plant Sci. 7:1049. doi: 10.3389/fpls.2016.01049

INTRODUCTION

Simple sequence repeats (SSRs), also known as microsatellites, are composed of 1–6 nucleotide motifs that are repeated in tandem and are widely and non-randomly distributed in 100–1000s of copies in the genomes of both monocots and dicots (Tautz and Renz, 1984; Toth et al., 2000; Mortimer et al., 2005; Lawson and Zhang, 2006; Hong et al., 2007). Microsatellites are found predominantly in heterochromatin regions, such as centromeric, peri-centromeric, and sub-distal regions of eukaryotic chromosomes (Yang et al., 2005; Li et al., 2010), and sex chromosomes in animals (Beckmann and Weber, 1992; Cuadrado and Jouve, 2011), usually associated with

the constitutive heterochromatin (Lohe et al., 1993; Pedersen et al., 1996). Variously non-random, taxon-specific patterns of SSR occurrence call for functional interpretations. Morgante et al. (2002) hypothesized that the relative frequency of microsatellites is higher in the single- or low-copy regions of the genome than in the repetitive regions. Mortimer et al. (2005) showed that SSRs are associated with and around transcribed sequences in *Arabidopsis*. Li et al. (2004) reviewed that SSRs in different positions in a gene can play important roles in regulating its expression and determining the function of its products. The accumulated evidence indicates that SSRs play an important role in chromatin organization, regulation of gene activity (Nagaki et al., 2004), recombination, DNA replication, the cell cycle, the mismatch DNA repair system (Li et al., 2002), and protein coding regions (Sonah et al., 2011).

The fluorescence *in situ* hybridization (FISH) technique has been used to localize one or more SSR loci on chromosomes. FISH has become a strategy for chromosome diagnosis and for investigating plant genome organization (Cuadrado and Schwarzacher, 1998; Cuadrado and Jouve, 2002). SSRs change rapidly during evolution, and thus display polymorphism at homologous sites between closely related species. Begum et al. (2009) showed that the rapid evolution of repetitive DNA sequences has resulted in species-specific repeat variants and the generation of novel repeat families. This characteristic has made SSRs useful as markers in comparative diversity analysis (Zhebentyayeva et al., 2003; Yu et al., 2010; Fang et al., 2013; Zhu et al., 2013) and genetics research (Nanda et al., 1991). SSRs are highly abundant within genomes, they can be widely dispersed or be confined to certain chromosomal regions, and they display a high degree of length polymorphism (Katti et al., 2001; Lawson and Zhang, 2006). Carmona et al. (2013) defined cytogenetic diversity in terms of the differences in abundance and distribution of microsatellites, and also found some specific and motif-dependent hybridization patterns. The repeats AG, AAG, ACT, and ATC presented different *in situ* hybridization patterns that provided cytogenetic landmarks for chromosome identification in barley, *Hordeum vulgare* ssp. *vulgare* (Carmona et al., 2013). Altogether, such variation could be used to determine evolutionary relationships between related species.

Brassica rapa belongs to the A genome species group in the Brassicaceae with $2n = 20$ chromosomes (Nagaharu, 1935), which had a monophyletic origin (Lysak et al., 2005; Cheng et al., 2013). *B. rapa* comprises several sub-species, such as non-heading Chinese cabbage (*B. rapa* ssp. *chinensis*), Chinese cabbage (*B. rapa* ssp. *pekinensis*; Koo et al., 2004), and turnip (*B. rapa* L. ssp. *Rapifera*; Snowdon, 2007). *B. rapa* ssp. *chinensis* is one of the most important leafy vegetable forms in *B. rapa*, and consists of five morphotypes (Pak-choi, Wu ta cai, Cai xin, Fen nie cai, and Tai cai; Viehoever et al., 1920[1]; Gladis and Hammer, 1992; Zheng et al., 2015). The phylogenetic relationships between some *B. rapa* ssp. *chinensis* morphotypes have been determined from morphological, ecological, and molecular data (Yu et al., 2010). In this paper, systematic research was performed to investigate the cytogenetic diversity between intra-specific forms of *B. rapa*.

The objective of the work presented here is to characterize the cytogenetic diversity of SSRs between morphotypes of *B. rapa* ssp. *chinensis* that represent a broad range of cytogenetic diversity. The available genome sequence of Chinese cabbage is an important and fundamental resource for understanding this species[2] (Wang et al., 2011), but the vast majority of heterochromatic regions remain essentially uncharacterized. The current estimates of SSR frequencies in many organisms differs from reality after comparisons with sequence databases (Hong et al., 2007; Cavagnaro et al., 2010; Gao et al., 2011). The distribution of SSRs in databases has been reported for the Chinese cabbage genome (Hong et al., 2007). In view of their ubiquity and functional importance, detailed information will be necessary to explore the comparative cytogenetics of SSRs in *B. rapa*.

Simple sequence repeats appear to be more abundant in non-coding regions than in coding regions of plant genomes (Hong et al., 2007; Cavagnaro et al., 2010). Tri-nucleotide repeats, the most abundant SSR types in many species (Gao et al., 2003; Shi et al., 2013), were found to be the most frequent in protein coding regions (Sonah et al., 2011). In addition to tri-nucleotide repeats, mono- and di-nucleotide repeats were also predominant in the *B. rapa* genome (Hong et al., 2007; Gao et al., 2011). In this study, we selected mono-, di-, and tri-nucleotide repeats for physical mapping on the chromosomes of *B. rapa*. FISH was performed with mono-, di-, and tri-nucleotides to detect the distributional profile of SSRs and to enhance our understanding of genome organization. The distributional characterization of SSRs revealed a range of cytogenetic diversity that could relate to genome organization, function, and evolutionary trends. We demonstrated that: (1) not all of the SSR-based probes produced FISH signals on all *B. rapa* ssp. *chinensis* chromosomes; (2) some SSR signal intensity did not show a relationship to the abundance in the genome database; (3) the distributional patterns of SSR signals depended on the SSR motif used and the species analyzed; and (4) differences in SSR abundance and density were shown within and between genomes.

MATERIALS AND METHODS

Plant Materials

Five morphotypes of *B. rapa* ssp. *chinensis*: Pak-choi (cv. NHCC002), Wu ta cai (cv. NHCC006), Cai xin (cv. NHCC008), Fen nie cai (cv. NHCC010), Tai cai (cv. NHCC015) were stored and cultivated at the Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in East China, Ministry of Horticulture, China. All of these morphotypes have distinct phenotypic characteristics in terms of leaf shape and size, number of leaves, number of outgrowing axially buds, and flowering time (**Figures 1a–e**).

Chromosome Preparation

Mitotic metaphase chromosome preparation from root tips followed the procedure described by Zheng et al. (2014).

[1]http://www.plantnames.unimelb.edu.au/Sorting/Brassica_rapa.html#praecox

[2]http://brassica.bbsrc.ac.uk/brassica_genome_sequencing_concept.htm



FIGURE 1 | The five morphotypes of *Brassica rapa* ssp. *chinensis*. (a) Pak-choi; **(b)** Wu ta cai; **(c)** Cai xin; **(d)** Fen nie cai; **(e)** Tai cai.

Seeds of all morphotypes were allowed to germinate on moist filter paper in Petri dishes at 25°C until each root was approximately 1.5 cm long. To increase the number of cells at metaphase, seedlings were treated with 2.0 μM 8-oxyquinoline at room temperature for 1.5–2.0 h. After washing three-times for 5 min each in distilled water, the seedlings were fixed in a fresh 3:1 (v/v) mixture of 100% ethanol:glacial acetic acid for 24 h, and preserved in 70% (v/v) ethanol. Root tips were digested with 4% (w/v) cellulase plus 2% (w/v) pectinase for approximately 30 min at 37°C, after which they were squashed in a drop of 45% (v/v) acetic acid. After removing the cover slip by freezing, each slide was air-dried in preparation for FISH.

Fluorescence *In situ* Hybridization (FISH)

Table 1 shows all of the synthetic oligonucleotides of 10–20 bp that were used as SSR probes. For mono-nucleotides, the rate of A or C repeats was representative of itself and T and G. The AC/AG motifs represent both themselves and the complementary sequences TG/TC. Two di-nucleotides (AT and GC) were not be used in FISH for their self-complementary structure (Cuadrado et al., 2008). Hybridization with tri-nucleotides (Jurka and Pethiyagoda, 1995), together with mono- and di-nucleotides, were performed on metaphase chromosomes of five representative cultivars of *B. rapa* ssp. *chinensis* morphotypes. The SSR sequences were synthesized by Life Technologies (Nanjing, China), and were labeled with digoxigenin-11-dUTP by random primer labeling followed the manufacturer's instructions (Roche). The reaction was performed with 2.0 μl of SSR in a 20 μl standard reaction by PCR for 3 h at 37°C, and the reaction was stopped at 65°C for 5 min. Probes were stored at −20°C prior to use in hybridizations.

Fluorescence *in situ* hybridization was performed as described by Zheng et al. (2014). The post-hybridization slide washing procedure was that of Heslop-Harrison (1991). Detection of digoxigenin was performed by incubating the slides in anti-digoxigenin-rhodamine (Roche) at 37°C for 1 h. The chromosomes were then counterstained with 2 μg/μl DAPI (Sigma). Re-probing was performed following the method of Cuadrado and Jouve (1994).

TABLE 1 | The simple sequence repeat (SSR) probes used in this study.

Nucleotide repeat type	SSR probes
Mono-nucleotides	$(A)_{10}$, $(C)_{10}$
Di-nucleotides	$(AC)_8$, $(AG)_{12}$
Tri-nucleotides	$(GCC)_5$, $(ACG)_5$, $(ACT)_5$, $(CAG)_5$, $(CAT)_5$, $(CAC)_5$, $(ATT)_5$, $(AAC)_5$, $(AAG)_5$, $(AGG)_5$, $(AGC)_5$, $(ATC)_5$

Image Acquisition and Analysis

Fluorescence *in situ* hybridization signals and images of stained chromosomes were captured using a chilled charge-coupled device (CCD) camera (Axiocam HR, Carl Zeiss, Germany), and images were pseudo-colored and processed using Axiovision software (Carl Zeiss). Detection signals and imaging acquisition were obtained by Zeiss Axio Imager A1 fluorescence microscope. For each SSR motif experiment, we analyzed at least 10 cells with distinct signals. The images from FITC and DAPI staining procedures were recorded separately using a cooled CCD camera. The exposure times depended on the intensity of the signals from each probe. The final images were prepared with Adobe Photoshop, version CS4.

RESULTS

We used 16 synthetic SSRs as probes for single- and double-target FISH. Differences were observed in the abundance and localization of motifs between the different *B. rapa* ssp. *chinensis* morphotypes, although a general distribution pattern emerged. In all five genomes, most motifs showed a higher density of signal at the centromeric or peri-centromeric regions.

Distribution of Mono- and Di-nucleotide SSRs in the Genomes of *B. rapa* ssp. *chinensis* Morphotypes

We did not detect visible signals for the mono-nucleotide repeats A and C on chromosomes of any target genome of *B. rapa* ssp. *chinensis*. The di-nucleotide probes, AC and AG, gave visible signals with different patterns among the five samples (**Figure 2**). Cai xin was the only morphotype in which we detected distinct signals on chromosomes from the two di-nucleotide probes. AG microsatellites showed weak hybridization signals with dispersed patterns in the genomes of Pak-choi and Fen nie cai (**Figures 2f,g**).

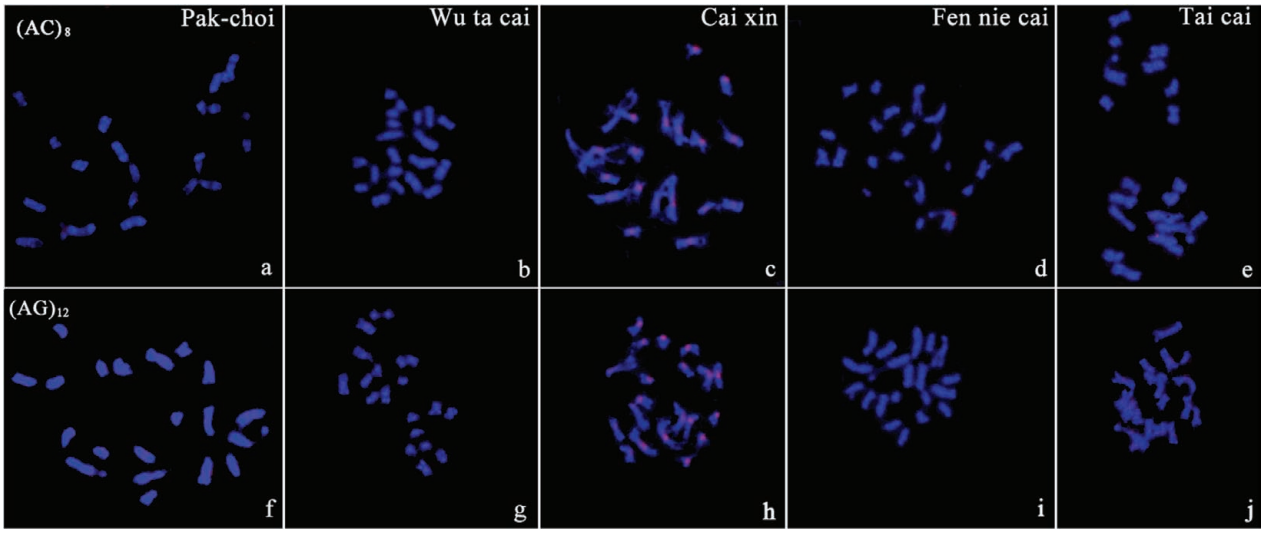


FIGURE 2 | Characterization of the di-nucleotide repeats AC and AG on chromosomes of *B. rapa* ssp. *chinensis* morphotypes by fluorescent *in situ* hybridization (FISH) using digoxigenin-labeled probes (detected with red rhodamine). (a–e) $(AC)_8$ in Pak-choi, Wu ta cai, Cai xin, Fen nie cai, Tai cai; **(f–j)** $(AG)_{12}$ in Pak-choi, Wu ta cai, Cai xin, Fen nie cai, Tai cai. Chromosomes were counterstained with DAPI.

Physical Characterization of Tri-nucleotide Repeats in *B. rapa* ssp. *chinensis* Morphotypes

Chromosomal Localization Tri-nucleotide Repeats in Pak-choi

Metaphase chromosomes of five *B. rapa* ssp. *chinensis* morphotypes were hybridized by re-probing preparations with tri-nucleotide repeat probes (**Supplementary Figures S1–S3**). All microsatellite motif probes gave *in situ* hybridization signals on metaphase chromosomes of Pak-choi. Well-defined hybridization signals were produced by ATT, CAC, CAT, AGG, AGC, and ATC probes, and these sequences showed characteristic, motif-dependent distribution patterns (**Figure 3**). The ATT and CAC probes revealed weak hybridization signals restricted to the centromeres of chromosomes A4 and A5, and the CAT repeat probe gave no signal (**Figures 3a–c**). In addition, no signal was detected on chromosome A6 after hybridization with CAC, CAT, and ATT repeat probes. The AGC microsatellites co-localized with ATC repeats on Pak-choi metaphase chromosomes with similar intensity, but showed less intensity than AGG repeats (**Figures 3d–f**). AGG, AGC, and ATC probes showed obvious differences in signal intensity on chromosome A2, AGG being the most intense. These three SSR clusters differed extensively on chromosome A9. AGC and ATC repeats were confined to centromeric regions of this chromosome (**Figures 3e,f**); however, AGG repeats comprised nearly the entire length of the short arm of A9 (**Figure 3d**).

Chromosomal Localization Tri-nucleotide Repeats in Wu ta cai

Differences in the presence/absence and intensity of the hybridization signals were observed on chromosomes of Wu ta cai after hybridization with tri-nucleotide repeat probes (**Figures 4a,b**). All tri-nucleotide repeats were near the centromere on some chromosomes. Polymorphic intercalary signals were observed in two regions of chromosome A2 after hybridizing with AGC and ATC clusters (**Figures 4i,k**). Hybridization signals of differing intensity were obtained, depending on the SSR motif probe. ATC and AGC, which were present on all chromosomes, showed obvious differences in intensity, the former being more intense. A weak signal on the long arm of chromosome A5, after hybridization with and AGC repeat probe, was observed only after increasing the exposure time of the CCD (compare **Figures 4h** and **4j**).

Chromosomal Localization Tri-nucleotide Repeats in Cai xin

Variation in intensity and location of the *in situ* signals were observed on metaphase chromosomes of Cai xin. The signals were confined to the centromere for all tri-nucleotide repeats with differing intensities. Polymorphism in terms of presence/absence was observed in the nucleolus organizing region (NOR). NOR signals were generated after hybridization with AG, CAC, ACT, GCC, and AGG repeats (**Figures 5b,c,e–g**). Visible centromeric signals were observed on chromosome A5 for AG, ACG, CAC, and ACT repeat probes (**Figures 5b–e**), but not for AC, GCC, and AGG repeats (**Figures 5a,f,g**).

Chromosomal Localization Tri-nucleotide Repeats in Fen nie cai

No specific clusters of SSR-specific signals were observed for tri-nucleotide repeats on chromosomes of Fen nie cai. Polymorphism in terms of presence/absence and intensity of the hybridization signals were observed for some SSRs. The absence signal was revealed from ACT and CAG clusters (**Supplementary Figures S1n,s**). However, signals of varying intensity were

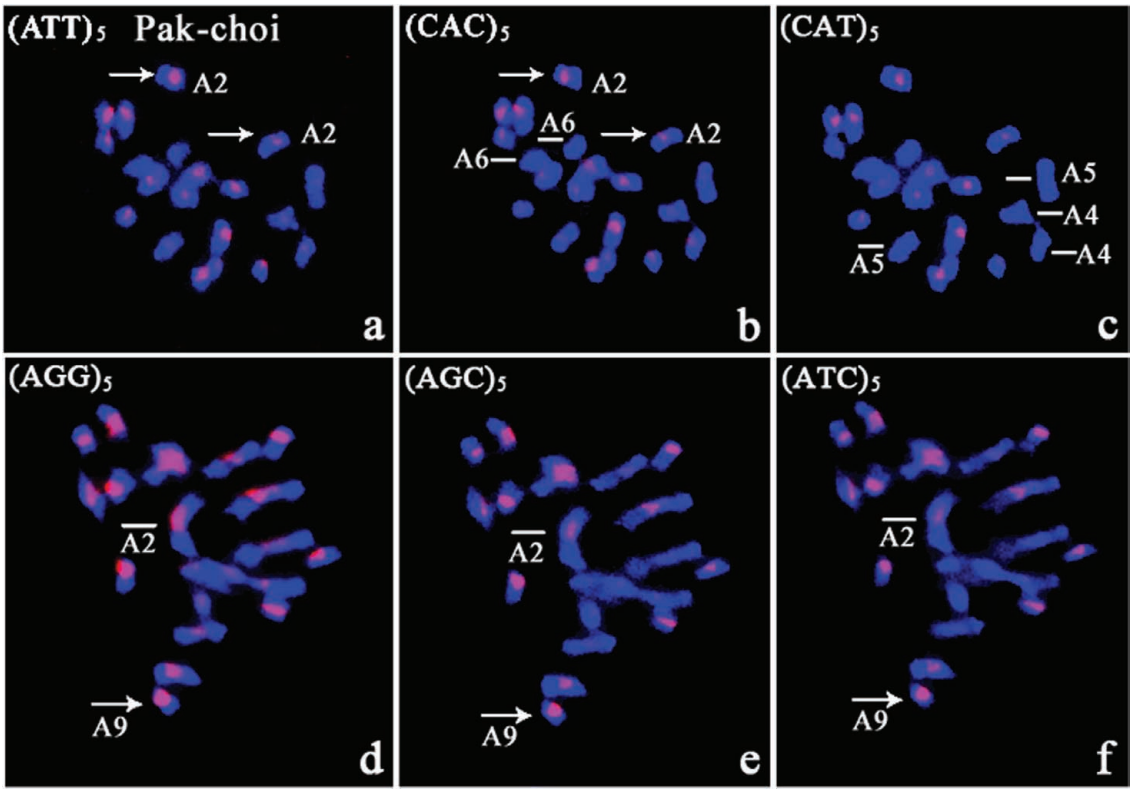


FIGURE 3 | FISH with CAC and ATT repeat probes on metaphase chromosomes of Pak-choi following *in situ* hybridization with digoxigenin-labeled probes (detected with red rhodamine) and DAPI counterstaining. (a–c) $(ATT)_5$, $(CAC)_5$, and $(CAT)_5$ in metaphase chromosomes of Pak choi; *arrows* indicate the different fluorescent sites from $(ATT)_5$ and $(CAC)_5$, and *lines* indicate no signal on chromosome A6 from $(CAC)_5$, chromosomes A4 and A5 from $(CAT)_5$; **(d–f)** $(AGG)_5$, $(AGC)_5$, and $(ATC)_5$ in metaphase chromosomes of Pak choi; *arrows* indicate the different fluorescent signal sites, and *lines* indicate different intensity fluorescent signals for the SSR loci.

observed on the homologous chromosomes after hybridization with AGG, AGC, and ATC repeat probes (**Figures 4c–e**).

Chromosomal Localization Tri-nucleotide Repeats in Tai cai

The most intense and rich patterns of *in situ* hybridization signals were produced by ACG, ACT, and ATT repeat probes on chromosomes of Tai cai, and were confined to the peri-centromeric regions. No specific clustered sites were observed for GCC and AAC repeats (**Supplementary Figures S1e** and **S3t**). AAT and CAC repeats were confined around centromeric regions. The ATT repeat probes gave more dispersed and intense centromeric signals, clustered in centromeric and peri-centromeric regions, than did the CAC repeats (**Figures 4f,g**). Variations in the number of signal sites were observed on chromosomes A2 and A6. An intercalary signal was observed on the long arms of chromosome A2 (compare **Figures 4l** with **4n**) and A6 (compare **Figures 4m** with **4o**) from hybridization with ATT repeats, in contrast to only one centromeric signal from CAC repeats (**Figures 4n,o**). No signals were observed from hybridization with a CAC repeat probe on chromosomes A4, A5, and the NOR, but the ATT repeat probe gave weak signals (**Figures 4f,g**).

DISCUSSION

Rogan et al. (2001) showed that base composition, probe length, and chromosomal location contribute to hybridization signal intensity in FISH. In the present study, all probes ranged from 15 to 20 bp in length. The hybridization patterns obtained depended on the base composition of the probes used and chromosomal location. The synthetic SSR probes were labeled by the random primer method, compared with end labeling, which improved the resolution of SSR loci in FISH (Bouilly et al., 2008). Different SSR probes of the same lengths gave different signal intensities, indirectly reflecting the influence of the target size and the copy numbers of the repeat sequences. The SSR probes gave specific hybridization patterns that provide cytogenetic landmarks for chromosome identification.

Different SSR Distribution Patterns on Chromosomes Depend on the Species Analyzed

Sonah et al. (2011) showed that mono-, di-, and tri-nucleotide repeats compose the major proportion of SSRs in plant genomes. For tri-nucleotides, A/T-rich repeats (e.g., AAC/GTT, AAG/CTT, and AAT/ATT) were predominant in dicot species. In the

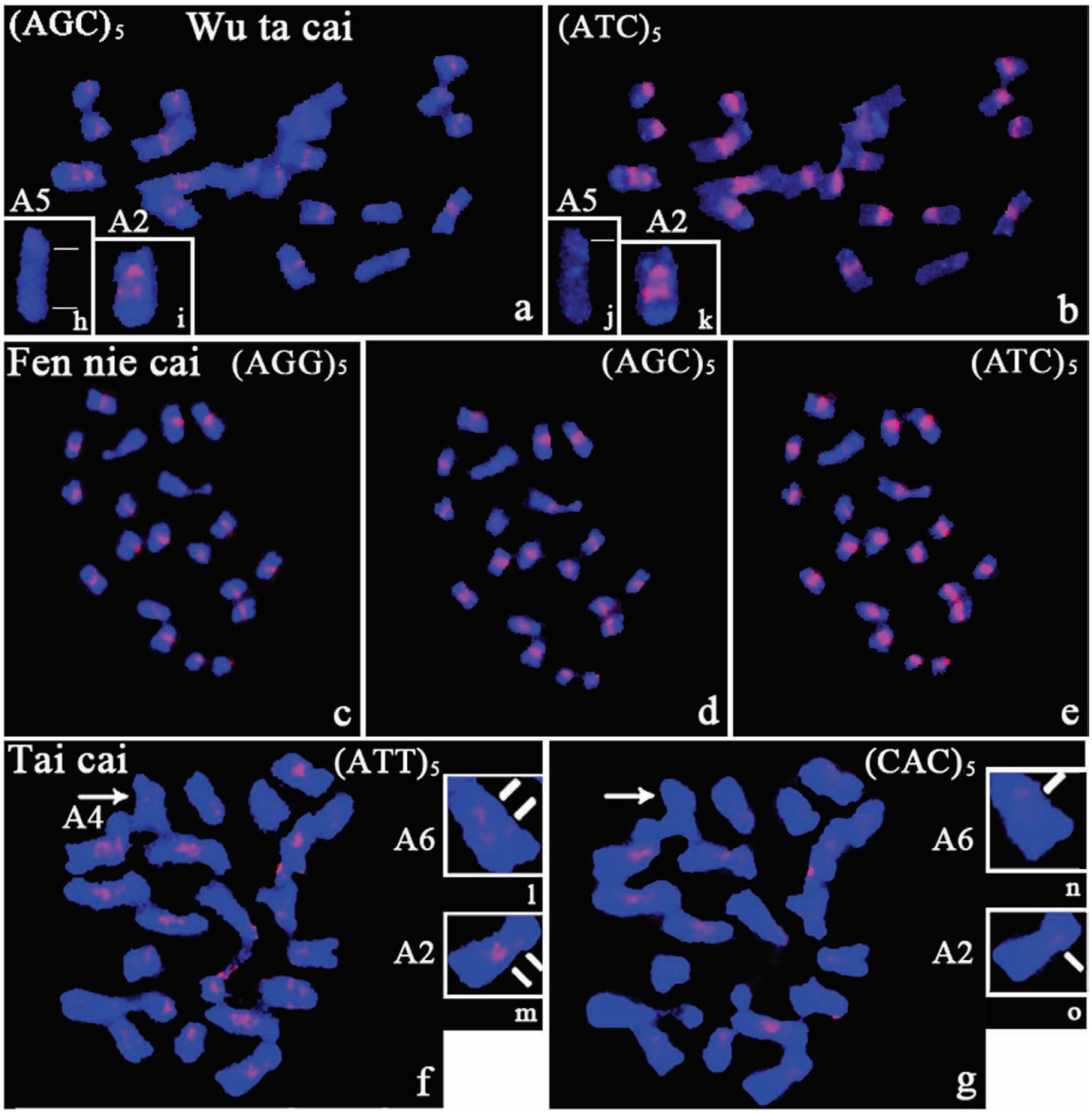


FIGURE 4 | Photomicrographs showing the distribution of several tri-nucleotide SSRs (AGG, AGC, and ATC) on metaphase chromosomes of Wu ta cai, Fen nie cai, and Tai cai. (a,b) in Wu ta cai; **(c–e)** in Fen nie cai; **(f,g)** in Tai cai; **(h–k)** variation in the number of signal sites in Wu ta cai; **(l–o)** variation in the number of signal sites in Tai cai; *lines* indicate the number of fluorescent signals in Wu ta cai and Tai cai; *arrows* indicate different intensity fluorescent signals. Chromosomes were counterstained with DAPI.

monocot barley, however, repeated AAT SSR motifs gave poor hybridization signals (Cuadrado and Jouve, 2007b). Morgante et al. (2002) found that GCC repeats accounted for half of the tri-nucleotide repeats in rice, whereas they were rare in dicots. Lawson and Zhang (2006) examined the most common mono- (A/T) and di-nucleotide (AT and AG) repeats in the *Arabidopsis* genome, and found that polyA/T repeats were predominant, while polyC/G repeats were rare (Sonah et al., 2011). The distribution of AC and AG repeats were linked to the euchromatic and heterochromatic genomic regions, respectively (Cuadrado and Jouve, 2007a). The di-nucleotide AG repeats were located on all chromosomes in *Dendrobium aphyllum* and *D. aggregatum* (Begum et al., 2009), but were exclusively concentrated at the centromeres in *Triticum* (Carmona et al., 2013). The most abundant repeat motifs were A (28.8%), AG (15.4%), AT (13.7%), and AAG (13.3%) clusters, reflecting the A/T rich nature of the *B. rapa* genome (Hong et al., 2007). In this study, nearly all the di- and tri-nucleotide SSRs were detected on metaphase chromosomes of Cai xin and Pak-choi, inferring that these two morphotypes have more types and increased abundance of SSRs

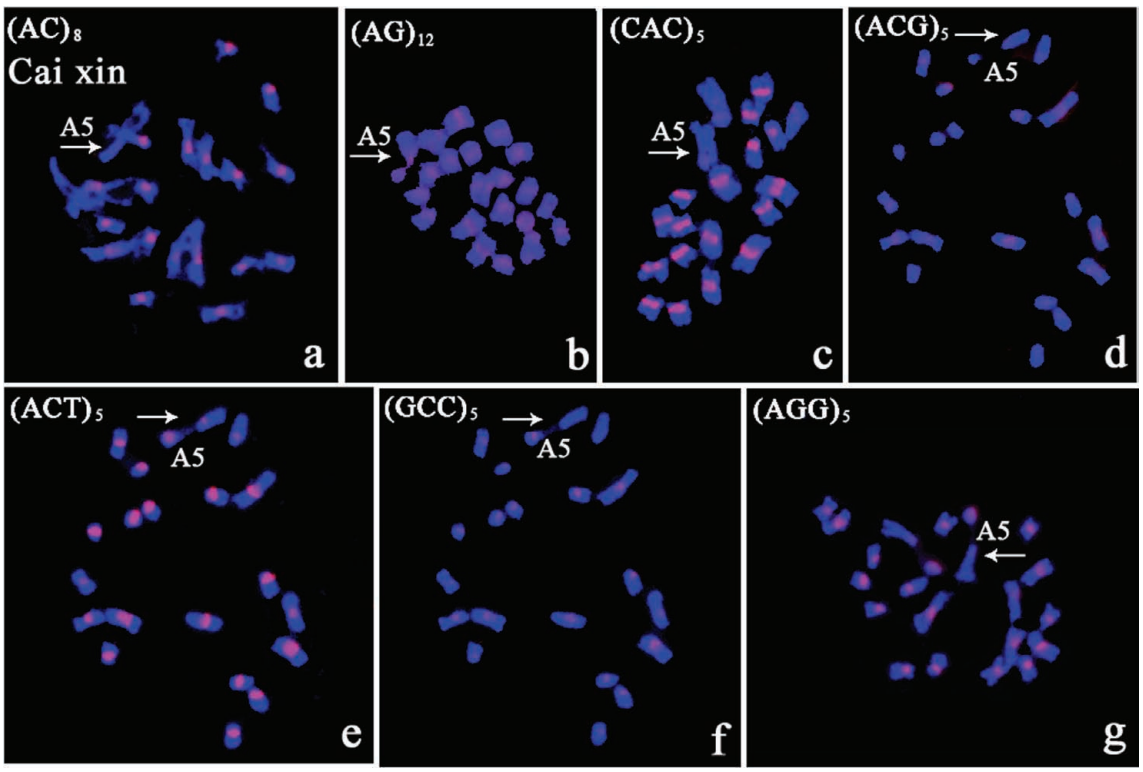


FIGURE 5 | The polymorphic characterization of the chromosome bearing the NOR in Cai xin following *in situ* hybridization with digoxigenin-labeled probes (detected with red rhodamine) and DAPI counterstaining. (a) $(AC)_8$; **(b)** $(AG)_{12}$; **(c)** $(CAC)_5$; **(d)** $(ACG)_5$; **(e)** $(ACT)_5$; **(f)** $(GCC)_5$; **(g)** $(AGG)_5$; *arrows* indicate chromosome A5; no signal was detected on the NOR in **(a**,**d)**; fluorescent signal on the NOR in **(c)**; weak centromeric signal in **(d)**; signals on the NOR and the centromere in **(b**,**e**,**g)**.

compared to the other three morphotypes (Wu ta cai, Fen nie cai, and Tai cai).

Various types of SSR motifs display taxon-specific patterns in the genomes of prokaryotes and eukaryotes (Toth et al., 2000). For example, the most intense hybridization signals were produced by the AGG, AAG, and AAC tri-nucleotide probes in barley (Cuadrado and Jouve, 2007b). The AAC clusters showed the same distribution patterns as AAG repeats in wheat (Cuadrado and Schwarzacher, 1998). AAG repeat units are major contributors to the genomes of dicots (Sonah et al., 2011); they are preferentially associated with peri-centromeric heterochromatin in *Hordeum* species (Carmona et al., 2013), and generally reflect the distribution of heterochromatin and the C-banding pattern in wheat (Cuadrado and Jouve, 1994). AAC repeats are organized in a more dispersed manner, with centromeric regions being largely excluded in chickpea and tomato (Gortner et al., 1998; Gindullis et al., 2001). Some repeats, such as CAG, CAC, and ACG, have specific hybridization sites restricted to the centromeres of metaphase chromosomes in barley (Cuadrado and Jouve, 2007b).

Despite the stable chromosome number and similarities in chromosome size and morphology, differences in numbers and distribution of SSR blocks have been observed among *B. rapa* ssp. *chinensis* morphotypes. The ATT repeat probe gave two signals with different intensities on chromosomes A2 and A6 of Tai cai, indicating that this SSR experienced chromosome-specific accumulation during evolution, as did AGG and ATC repeats in Wu ta cai. In addition, several SSRs (ATT, CAC, AGG, AGC, and ATC) also showed chromosome-specific signals on marker chromosomes of Pak-choi. These three morphotypes may have resulted from chromosomal recombination during speciation and development.

The Different Distribution Patterns on Chromosomes Depend on the SSR Probes Used

A negative correlation has been observed between repeat numbers and the total length of repeat units in both monocots and dicots (Cavagnaro et al., 2010; Gao et al., 2011; Sonah et al., 2011). Mono-nucleotide repeats of A are the most abundant repeats among all SSRs analyzed (Gao et al., 2011); however, no visible signal was generated, partly implying their dispersed distribution along chromosomes, or may be related to the centromeric function (Cuadrado and Jouve, 2007a). In this work, the di-nucleotide repeats AC and AG only showed distinct *in situ* hybridization signals on chromosomes of Cai xin, although there were weak signals in Pak-choi and Fen nie cai for AG repeats. In addition, three tri-nucleotide repeat probes (AAG, AGG, and AGC) gave distinct signals on all analyzed genomes. Di-nucleotide AG and tri-nucleotide AAG repeats are

relatively abundant in the *B. rapa* genome (Hong et al., 2007). AG-rich microsatellites might be the most prevalent SSRs that cluster around centromeric regions in *B. rapa* ssp. *chinensis* morphotypes.

We found that some tri-nucleotide repeats distinctly co-localized within the same genome, although the signals differed in intensity, which could be explained by their intermixed structure at the given loci. Some SSR probes gave weak or dispersed signals at the same physical position, which could possibly be due to closely linked blocks of repeated sequences. Here, we show that individual SSRs vary widely in their relative proportions at the chromosome level, and that the distribution of SSRs along the chromosomes is non-random. The characteristic patterns of SSR distribution show that centromeric regions are more densely populated than the central regions (Rogan et al., 2001). The similar distribution patterns of some SSR motifs indicate that long stretches of different SSRs are of functional importance, and could possibly represent an ancient component of plant genomes (Cuadrado et al., 2000; Cuadrado and Jouve, 2007b). These results are supported by evidence showing that microsatellites display relatively uniform coverage in the genome, and that there are taxon-specific distribution patterns among *B. rapa* ssp. *chinensis* morphotypes. Our results suggest that SSR repeats are the major component of the satellite DNA fraction, and show evolutionary conservation among *B. rapa* ssp. *chinensis* morphotypes.

Relationship between SSRs and Centromeres

In most higher eukaryotic organisms, chromosomal centromeres are composed of long arrays of satellite repeat sequences and retro-transposons (Henikoff et al., 2001; Jiang et al., 2003). The centromeric and distal regions play important roles during mitosis and meiosis (Li et al., 2002; Hong et al., 2006), and highly divergent sequences are present in the peri-centromeric regions (Wang et al., 2012). The repetitive DNA sequences frequently form clusters within heterochromatin blocks, which are predominantly concentrated at peri-centromeric regions and have been detected in plants with small and compact genomes (Cuadrado and Jouve, 2010; Falistocco and Marconi, 2013). In the present work, most *in situ* SSR signals were confined to centromeric or adjacent regions on *B. rapa* ssp. *chinensis* chromosomes. This distributional characterization may be related to their effect on DNA replication, chromatin organization, and the cell cycle.

Evolutionary Trends of SSRs among *B. rapa* ssp. *chinensis* Genomes

The relationship between microsatellites and chromosomal evolution has not been clearly documented. The frequency of repeats decreases exponentially with their length, type, and number of SSR motif repeats. This characterization appears to be more conservative in coding than non-coding sequences (Cavagnaro et al., 2010; Sonah et al., 2011; Shi et al., 2013), and less pronounced for di-nucleotides compared to longer repeat types (Cavagnaro et al., 2010). The trends for various repeat types are similar between different chromosomes within the same genome, but the density of repeats may vary between different chromosomes in the same species (Schafer et al., 1986; Katti et al., 2001). Sonah et al. (2011) found species-specific accumulation of particular motif repeats. Variation is present between related species in terms of the abundance and chromosomal distribution of SSR clusters among morphotypes. Schmidt and Heslop-Harrison (1996) demonstrated that microsatellites, representing a substantial fraction of the genome, showed chromosome-specific amplification in plants. High levels of polymorphism and heterozygosity between homologs, in terms of the distribution of AAG and AAC repeats, was shown for out-breeding species in the *Secale strictum* species complex (Cuadrado and Jouve, 2002).

Our results suggest that SSR sequences are more predisposed to being amplified or deleted as a result of independent events. The balance among SSR sequences generated by strand-slippage replication, or recombination and repair mechanisms, cannot be the only explanation for the observed differences in their chromosomal distributions. Another possible explanation would involve selection pressure or mutation (Li et al., 2002). The different chromosomal positions of SSRs involved in the regulation of gene expression (Lawson and Zhang, 2006; Gao et al., 2011), could indicate their underestimated roles in genome evolution (Cuadrado et al., 2008).

The microsatellite sequences analyzed here showed similar chromosome distribution polymorphism patterns, inferring that these SSR loci may result from convergent evolution. However, they differed in intensity or position, indicating that microsatellite repeats can contract or expand over a very short evolutionary time frame (Iwata et al., 2013). The wide distribution of SSRs, and the fact that their positions are restricted to chromosomal centromeres, as revealed by FISH, suggested a general model for the parallel chromosome evolution of repeat-rich heterochromatin in *B. rapa* ssp. *chinensis*.

Carmona et al. (2013) suggested that changes in the amount and distribution of tandem repetitive DNA sequences are major driving forces of genome evolution and speciation. The different regions are thought to undergo different selection pressures (Morgante et al., 2002), which might account for different motif preferences and frequencies among chromosomes. The evolutionary dynamics of microsatellites is generally consistent with plant divergence and evolution (Shi et al., 2013), and the distribution of microsatellites is related to the history of genome evolution and selective constraints (Morgante et al., 2002). The variation in SSRs at the chromosome level may be the result of adaptive divergence, or selection resulting from the stress response among species and populations (Cavagnaro et al., 2010; Gao et al., 2011; Carmona et al., 2013). Whether SSRs are under selection or are neutral as has been reported (Ellegren, 2004), and can be used for exploring the dynamics of the evolutionary process (Santos et al., 2010), will require study.

AUTHOR CONTRIBUTIONS

S-nZ and J-sZ: Designed the experiment. X-lH and GB: Provided the materials. J-sZ and C-zS: Performed the experiment. J-sZ: Analyzed the data and wrote the manuscript.

FUNDING

This work was supported by Science and Technology Pillar Program of Jiangsu Province (BE2013429), the Agricultural science and technology independent innovation funds of Jiangsu Province [CX(13)2006], the National Fund of Hebei Province, China (Project No. C2015407058) and Scientific Research Project of Hebei Province China (Project No. QN2016110).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01049

FIGURE S1 | Photomicrographs showing the distribution of the tri-nucleotide repeats $(GCC)_5$, $(ACG)_5$, $(ACT)_5$, $(CAG)_5$ on metaphase chromosomes of five *B. rapa* ssp. *chinensis* morphotypes after *in situ* hybridization with digoxigenin-labeled probes (detected with red rhodamine) and DAPI counterstaining. (a,f,k,p) in Pak-choi; **(b,g,l,q)** in Wu ta cai; **(c,h,m,r)** in Cai xin; **(d,i,n,s)** in Fen nie cai; **(e,j,o,t)** in Tai cai.

FIGURE S2 | Photomicrographs showing the distribution of $(AAG)_5$, $(AGG)_5$, $(AGC)_5$, and $(ATC)_5$ repeats on metaphase chromosomes of five *B. rapa* ssp. *chinensis* morphotypes after *in situ* hybridization with digoxigenin-labeled probes (detected with red rhodamine) and DAPI counterstaining. (a,f,k,p) in Pak-choi; **(b,g,l,q)** in Wu ta cai; **(c,h,m,r)** in Cai xin; **(d,i,n,s)** in Fen nie cai; **(e,j,o,t)** in Tai cai.

FIGURE S3 | Photomicrographs showing the distribution of $(CAT)_5$, $(CAC)_5$, $(ATT)_5$, and $(AAC)_5$ repeats on metaphase chromosomes of five *B. rapa* ssp. *chinensis* morphotypes after *in situ* hybridization with digoxigenin-labeled probes (detected with red rhodamine) and DAPI counterstaining. (a,f,k,p) in Pak-choi; **(b,g,l,q)** in Wu ta cai; **(c,h,m,r)** in Cai xin; **(d,i,n,s)** in Fen nie cai; **(e,j,o,t)** in Tai cai.

REFERENCES

Beckmann, J. S., and Weber, J. L. (1992). Survey of human and rat microsatellites. *Genomics* 12, 627–631. doi: 10.1016/0888-7543(92)90285-Z

Begum, R., Alam, S. S., Menzel, G., and Schmidt, T. (2009). Comparative molecular cytogenetics of major repetitive sequence families of three *Dendrobium* species (Orchidaceae) from Bangladesh. *Ann. Bot.* 104, 863–872. doi: 10.1093/aob/mcp178

Bouilly, K., Chaves, R., Leitao, A., Benabdelmouna, A., and Guedes-Pinto, H. (2008). Chromosomal organization of simple sequence repeats in the *Pacific oyster* (*Crassostrea gigas*): (GGAT)4, (GT)7 and (TA)10 chromosome patterns. *J. Genet.* 87, 119–125. doi: 10.1007/s12041-008-0018-2

Carmona, A., Friero, E., de Bustos, A., Jouve, N., and Cuadrado, A. (2013). Cytogenetic diversity of SSR motifs within and between *Hordeum* species carrying the H genome: *H. vulgare* L. and *H. bulbosum* L. *Theor. Appl. Genet.* 126, 949–961. doi: 10.1007/s00122-012-2028-y

Cavagnaro, P. F., Senalik, D. A., Yang, L. M., Simon, P. W., Harkins, T. T., Kodira, C. D., et al. (2010). Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics* 11:569. doi: 10.1186/1471-2164-11-569

Cheng, F., Mandáková, T., Wu, J., Xie, Q., Lysak, M. A., and Wang, X. W. (2013). Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell* 25, 1541–1554. doi: 10.1105/tpc.113.110486

Cuadrado, A., Cardoso, M., and Jouve, N. (2008). Physical organisation of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. *Cytogenet. Genome Res.* 120, 210–219. doi: 10.1159/000121069

Cuadrado, A., and Jouve, N. (1994). Mapping and organization of highly-repeated DNA sequences by means of simultaneous and sequential FISH and C-banding in 6x-triticale. *Chromosome Res.* 2, 331–338. doi: 10.1007/BF01552727

Cuadrado, A., and Jouve, N. (2002). Evolutionary trends of different repetitive DNA sequences during speciation in the genus Secale. *J. Hered.* 93, 339–345. doi: 10.1093/jhered/93.5.339

Cuadrado, A., and Jouve, N. (2007a). Similarities in the chromosomal distribution of AG and AC repeats within and between *Drosophila*, human and barley chromosomes. *Cytogenet. Genome Res.* 119, 91–99. doi: 10.1159/000109624

Cuadrado, A., and Jouve, N. (2007b). The nonrandom distribution of long clusters of all possible classes of tri-nucleotide repeats in barley chromosomes. *Chromosome Res.* 15, 711–720. doi: 10.1007/s10577-007-1156-8

Cuadrado, A., and Jouve, N. (2010). Chromosomal detection of simple sequence repeats (SSRs) using nondenaturing FISH (ND-FISH). *Chromosoma* 119, 495–503. doi: 10.1007/s00412-010-0273-x

Cuadrado, A., and Jouve, N. (2011). Novel simple sequence repeats (SSRs) detected by ND-FISH in heterochromatin of *Drosophila melanogaster*. *BMC Genomics* 12:205. doi: 10.1186/1471-2164-12-205

Cuadrado, A., and Schwarzacher, T. (1998). The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma* 107, 587–594. doi: 10.1007/s004120050345

Cuadrado, A., Schwarzacher, T., and Jouve, N. (2000). Identification of different chromatin classes in wheat using in situ hybridization with simple sequence repeat oligonucleotides. *Theor. Appl. Genet.* 101, 711–717. doi: 10.1007/s001220051535

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. *Nat. Rev. Genet.* 5, 435–445. doi: 10.1038/nrg1348

Falistocco, E., and Marconi, G. (2013). Cytogenetic characterization by in situ hybridization techniques and molecular analysis of 5S rRNA genes of the European hazelnut (*Corylus avellana*). *Genome* 56, 155–159. doi: 10.1139/gen-2013-0045

Fang, P., Chen, F. B., Yao, Q. L., Yang, K. C., Zhou, G. F., Fan, Y. H., et al. (2013). Analysis of genetic diversity in the tuber mustard (*Brassica juncea* var. tumida Tsen et Lee) in the Yangtze river basin of China. *Genet. Resour. Crop Evol.* 60, 129–143. doi: 10.1007/s10722-012-9821-y

Gao, C. H., Tang, Z. L., Yin, J. M., An, Z. S., Fu, D. H., and Li, J. N. (2011). Characterization and comparison of gene-based simple sequence repeats across *Brassica* species. *Mol. Genet. Genomics* 286, 161–170. doi: 10.1007/s00438-011-0636-x

Gao, L. F., Tang, J. F., Li, H. W., and Jia, J. Z. (2003). Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Mol. Breed.* 12, 245–261. doi: 10.1023/A:1026346121217

Gindullis, F., Desel, C., Galasso, I., and Schmidt, T. (2001). The large-scale organization of the centromeric region in Beta species. *Genome Res.* 11, 253–265. doi: 10.1101/gr.162301

Gladis, T., and Hammer, K. (1992). The Gatersleben Brassica collection–*Brassica juncea, B. napus, B. nigra and B. rapa* (Germ., Engl. summary). *Feddes Repert.* 103, 469–507. doi: 10.1002/fedr.19921030704

Gortner, G., Nenno, M., Weising, K., Zink, D., Nagl, W., and Kahl, G. (1998). Chromosomal localization and distribution of simple sequence repeats and the *Arabidopsis*-type telomere sequence in the genome of *Cicer arietinum* L. *Chromosome Res.* 6, 97–104. doi: 10.1023/A:1009282828236

Henikoff, S., Ahmad, K., and Malik, H. S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293, 1098–1102. doi: 10.1126/science.1062939

Heslop-Harrison, J. S. (1991). The molecular cytogenetics of palnts. *J. Cell Sci.* 100, 15–21.

Hong, C. P., Piao, Z. Y., Kang, T. W., Batley, J., Yang, T. J., Hur, Y. K., et al. (2007). Genomic distribution of simple sequence repeats in *Brassica rapa*. *Mol. Cells* 23, 349–356.

Hong, C. P., Plaha, P., Koo, D. H., Yang, T. J., Choi, S. R., Lee, Y. K., et al. (2006). A survey of the *Brassica rapa* genome by BAC-End sequence analysis and comparison with *Arabidopsis thaliana*. *Mol. Cells* 22, 300–307.

Iwata, A., Tek, A. L., Richard, M. M. S., Abernathy, B., Fonseca, A., Schmutz, J., et al. (2013). Identification and characterization of functional centromeres of the common bean. *Plant J.* 76, 47–60. doi: 10.1111/tpj.12269

Jiang, J. M., Birchler, J. A., Parrott, W. A., and Dawe, R. K. (2003). A molecular view of plant centromeres. *Trends Plant Sci.* 8, 570–575 doi: 10.1016/j.tplants.2003.10.011

Jurka, J., and Pethiyagoda, C. (1995). Simple repetitive DNA-sequeences from primates-compilation and analysis. *J. Mol. Evol.* 40, 120–126. doi: 10.1007/BF00167107

Katti, M. V., Ranjekar, P. K., and Gupta, V. S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Mol. Biol. Evol.* 18, 1161–1167. doi: 10.1093/oxfordjournals.molbev.a003903

Koo, D. H., Plaha, P., Lim, Y. P., Hur, Y., and Bang, J. W. (2004). A high-resolution karyotype of *Brassica rapa* ssp. pekinensis revealed by pachytene analysis and multicolor fluorescence in situ hybridization. *Theor. Appl. Genet.* 109, 1346–1352. doi: 10.1007/s00122-004-1771-0

Lawson, M. J., and Zhang, L. Q. (2006). Distinct patterns of SSR distribution in the *Arabidopsis thaliana* and rice genomes. *Genome Biol.* 7:R14. doi: 10.1186/gb-2006-7-2-r14

Li, C., Huang, B., Yan, X. H., Wang, L. J., Yang, Q., and Wei, W. H. (2010). Cloning, identification and characterization of a repetitive sequence flanking telomere and homologous to canrep in *Brassica napus*. *Bot. Stud.* 51, 421–430.

Li, Y., Korol, A. B., Fahima, T., Beiles, A., and Nevo, E. (2004). Microsatellites within genes: structure, function and evolution. *Mol. Biol. Evol.* 21, 991–1007 doi: 10.1093/molbev/msh073

Li, Y. C., Korol, A. B., Fahima, T., Beiles, A., and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol. Ecol.* 11, 2453–2465. doi: 10.1046/j.1365-294X.2002.01643.x

Lohe, A. R., Hilliker, A. J., and Roberts, P. A. (1993). Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics* 134, 1149–1174. doi: 10.1016/0168-9525(93)90135-5

Lysak, M. A., Koch, M. A., Pecinka, A., and Schubert, I. (2005). Chromosome triplication found across the tribe Brassiceae. *Genome Res.* 15, 516–525. doi: 10.1101/gr.3531105

Morgante, M., Hanafey, M., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. *Nat. Genet.* 30, 194–200. doi: 10.1038/ng822

Mortimer, J. C., Batley, J., Love, C. G., Logan, E., and Edwards, D. (2005). Simple sequence repeat (SSR) and GC distribution in the *Arabidopsis thaliana* genome. *J. Plant Biotechnol.* 7, 17–25.

Nagaharu, U. (1935). Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiarmode of fertilization. *Jpn. J. Bot.* 7, 389–452.

Nagaki, K., Cheng, Z. K., Ouyang, S., Talbert, P. B., Kim, M., Jones, K. M., et al. (2004). Sequencing of a rice centromere uncovers active genes. *Nat. Genet.* 36, 138–145. doi: 10.1038/ng1289

Nanda, I., Zischler, H., Epplen, C., Guttenbach, M., and Schmid, M. (1991). Chromosomal organization of simple repeated DNA-sequeences used for DNA fingerprinting. *Electrophoresis* 12, 193–203. doi: 10.1002/elps.1150120216

Pedersen, C., Rasmussen, S. K., and Linde-Laursen, I. (1996). Genome and chromosome identification in cultivated barley and related species of the Triticeae (Poaceae) by in situ hybridization with the GAA-satellite sequence. *Genome* 39, 93–104. doi: 10.1139/g96-013

Rogan, P. K., Cazcarro, P. M., and Knoll, J. H. M. (2001). Sequence-based design of single-copy genomic DNA probes for fluorescence in situ hybridization. *Genome Res.* 11, 1086–1094. doi: 10.1101/gr.171701

Santos, J., Serra, L., Sole, E., and Pascual, M. (2010). FISH mapping of microsatellite loci from *Drosophila subobscura* and its comparison to related species. *Chromosome Res.* 18, 213–226. doi: 10.1007/s10577-010-9112-4

Schafer, R., Ali, S., and Epplen, J. T. (1986). The organization of the evolutionarily conserved GATA GACA repeats in the mouse genome. *Chromosoma* 93, 502–510. doi: 10.1007/BF00386791

Schmidt, T., and Heslop-Harrison, J. S. (1996). The physical and genomic organization of microsatellites in sugar beet. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8761–8765. doi: 10.1073/pnas.93.16.8761

Shi, J. Q., Huang, S. M., Fu, D. H., Yu, J. Y., Wang, X. F., Hua, W., et al. (2013). Evolutionary dynamics of microsatellite distribution in plants: insight from the comparison of sequenced *Brassica*, *Arabidopsis* and other angiosperm species. *PLoS ONE* 8:e59988. doi: 10.1371/journal.pone.0059988

Snowdon, R. J. (2007). Cytogenetics and genome analysis in *Brassica* crops. *Chromosome Res.* 15, 85–95. doi: 10.1007/s10577-006-1105-y

Sonah, H., Deshmukh, R. K., Sharma, A., Singh, V. P., Gupta, D. K., Gacche, R. N., et al. (2011). Genome-wide distribution and organization of microsatellites in plants: an insight into marker development in *Brachypodium*. *PLoS ONE* 6:e21298. doi: 10.1371/journal.pone.0021298

Tautz, D., and Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* 12, 4127–4138. doi: 10.1093/nar/12.10.4127

Toth, G., Gaspari, Z., and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Res.* 10, 967–981. doi: 10.1101/gr.10.7.967

Viehoever, A. (1920). Studies in mustard seeds and substitutes: I. Chinese colza (*Brassica campestris* chinoleifera Viehoever). *J. Agric. Res.* 20, 117–140.

Wang, K., Zhang, W. P., Gao, Y. J., Zhang, Z. X., Zheng, D. W., Zhou, B. L., et al. (2012). Localization of high level of sequence conservation and divergence regions in cotton. *Theor. Appl. Genet.* 124, 1173–1182. doi: 10.1007/s00122-011-1777-3

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Yang, T. J., Lee, S., Chang, S. B., Yu, Y., Jong, H., and Wing, R. A. (2005). In-depth sequence analysis of the tomato chromosome 12 centromeric region: identification of a large CAA block and characterization of pericentromere retrotranposons. *Chromosoma* 114, 103–117. doi: 10.1007/s00412-005-0342-8

Yu, S. C., Zhang, F. L., Wang, X. Y., Zhao, X. Y., Zhang, D. S., Yu, Y. J., et al. (2010). Genetic diversity and marker-trait associations in a collection of Pak-choi (*Brassica rapa* L. ssp *Chinensis* Makino) accessions. *Genes Genomics* 32, 419–428. doi: 10.1007/s13258-010-0033-6

Zhebentyayeva, T. N., Reighard, G. L., Gorina, V. M., and Abbott, A. G. (2003). Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm. *Theor. Appl. Genet.* 106, 435–444.

Zheng, J. S., Sun, C. Z., Xiao D., Zhang, S. N., Bonnema G., and Hou, X. L. (2015). Karyotype variation and conservation in morphotypes of non-heading Chinese cabbage. *Plant Syst. Evol.* 301, 1781–1791. doi: 10.1007/s00606-014-1177-7

Zheng, J. S., Sun, C. Z., Zhang, S. N., and Hou, X. L. (2014). Karyotype of mitotic metaphase and meiotic diakinesis in non-heading Chinese cabbage. *Plant Syst. Evol.* 300, 295–302. doi: 10.1007/s00606-013-0882-y

Zhu, X. C., Raman, H., Wu, H. W., Lemerle, D., Burrows, G. E., and Stanton, R. (2013). Development of SSR markers for genetic analysis of rilverleaf nightshade (*Solanum elaeagnifolium*) and related rpecies. *Plant Mol. Biol. Rep.* 31, 248–254. doi: 10.1007/s11105-012-0473-z

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer ST and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 Zheng, Sun, Zhang, Hou and Bonnema. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 October 2015
doi: 10.3389/fpls.2015.00836

Genome-wide DNA methylation profiling by modified reduced representation bisulfite sequencing in *Brassica rapa* suggests that epigenetic modifications play a key role in polyploid genome evolution

Xun Chen[1], Xianhong Ge[1], Jing Wang[1], Chen Tan[1], Graham J. King[2] and Kede Liu[1]*

[1] National Key Laboratory of Crop Genetic Improvement, National Center of Oil Crop Improvement (Wuhan), College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [2] Southern Cross Plant Science, Southern Cross University, Lismore, NSW, Australia

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Maoteng Li,
Huazhong University of Science and Technology, China
Seonghee Lee,
University of Florida, USA

***Correspondence:**
Xianhong Ge,
National Key Laboratory of Crop Genetic Improvement, National Center of Oil Crop Improvement (Wuhan), College of Plant Science and Technology, Huazhong Agricultural University, Hongshan, Shizishan Street No. 1, Wuhan, Hubei 430070, China
gexianhong@mail.hzau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *23 July 2015*
Accepted: *23 September 2015*
Published: *09 October 2015*

Citation:
Chen X, Ge X, Wang J, Tan C, King GJ and Liu K (2015) Genome-wide DNA methylation profiling by modified reduced representation bisulfite sequencing in Brassica rapa suggests that epigenetic modifications play a key role in polyploid genome evolution. Front. Plant Sci. 6:836. doi: 10.3389/fpls.2015.00836

Brassica rapa includes some of the most important vegetables worldwide as well as oilseed crops. The complete annotated genome sequence confirmed its paleohexaploid origins and provides opportunities for exploring the detailed process of polyploid genome evolution. We generated a genome-wide DNA methylation profile for *B. rapa* using a modified reduced representation bisulfite sequencing (RRBS) method. This sampling represented 2.24% of all CG loci (2.5×10^5), 2.16% CHG (2.7×10^5), and 1.68% CHH loci (1.05×10^5) (where H = A, T, or C). Our sampling of DNA methylation in *B. rapa* indicated that 52.4% of CG sites were present as ^{5m}CG, with 31.8% of CHG and 8.3% of CHH. It was found that genic regions of single copy genes had significantly higher methylation compared to those of two or three copy genes. Differences in degree of genic DNA methylation were observed in a hierarchical relationship corresponding to the relative age of the three ancestral subgenomes, primarily accounted by single-copy genes. RNA-seq analysis revealed that overall the level of transcription was negatively correlated with mean gene methylation content and depended on copy number or was associated with the different subgenomes. These results provide new insights into the role epigenetic variation plays in polyploid genome evolution, and suggest an alternative mechanism for duplicate gene loss.

Keywords: ***Brassica rapa,*** **DNA methylation, genome evolution, modified RRBS, polyploid**

Introduction

Polyploidy, where more than two complete sets of chromosomes reside within the same nucleus, is both pervasive and ancient in most eukaryotic lineages and also is particularly prevalent in plants (Jiao et al., 2011; Proost et al., 2011). Polyploidization results in gene duplication, redundancy, and increased genome size, following which a dynamic polyploid genome will experience extensive and rapid genome restructuring, genome downsizing and ultimately genetic "diploidization" at many loci (Soltis and Soltis, 1999; Wolfe, 2001). The mechanism of diploidization remains a mystery,

although the loss of duplicated copies from the genomes of ancient polyploid species known as "fractionation" has been considered a major force for plant genome evolution (Langham et al., 2004). Interestingly, gene losses within different genomes are usually unequal, with one of the duplicated genomes consistently losing significantly more genes than the others. This bias in gene loss was first observed in *Arabidopsis* (Thomas et al., 2006) and more recently in maize (Woodhouse et al., 2010; Schnable et al., 2011) and *Brassica rapa* (Wang et al., 2011; Tang et al., 2012), and is probably a general characteristic of major eukaryote lineages where paleopolyploidy has been involved (Sankoff et al., 2010).

In maize, a monocotyledon species that experienced tetraploidy 5–12 million years ago, biased gene losses between two complete subgenomes are both ancient and ongoing among diverse inbreds, primarily resulting from a mechanism of short deletions (Woodhouse et al., 2010). In particular, genes from the genome that has experienced less gene loss are expressed at a higher level than those from the other subgenome. This suggests that the bias in gene loss between the subgenomes may be the result of selection against loss of the genes responsible for the majority of the expression within a duplicated gene pair (Schnable and Freeling, 2011; Schnable et al., 2011). A corollary of this is that genes with lower expression levels are more readily deleted, since their removal is less likely to lower fitness, and so they escape purifying selection. Differential epigenetic marking of the genomes within an allopolyploid has been suggested as a mechanism underlying differential expression of genes retained in different subgenomes (Schnable et al., 2011; Diez et al., 2014).

Cytosine methylation is a major epigenetic mark and plays an important role in chromatin conformation, in silencing different types of repetitive sequence and in regulating transcription (Bird, 2002). In plants, methylated cytosine residues are observed at cytosine bases in all sequence contexts, including symmetric CG and CHG (where H = A, T, or C) and asymmetric CHH (Henderson and Jacobsen, 2007). Methylated cytosine (^{5m}C) is especially pervasive in intergenic regions but also within protein-coding regions, where it is typically limited to the CG context (Cokus et al., 2008; Lister et al., 2008). *De novo* methylation in plants is catalyzed by DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) but maintained by different pathways, with ^{5m}CG maintained by DNA METHYLTRANSFERASE1 (MET1), ^{5m} CHG by CHROMOMETHYLASE3 (CMT3), and asymmetric ^{5m} CHH dependent on persistent *de novo* methylation by DRM2 (Chan et al., 2005; Law and Jacobsen, 2010). Small RNA (sRNA)–mediated DNA methylation (RdDM) can target methylation of transposable elements (TEs) in many eukaryotic lineages and contribute to limiting TE proliferation (Almeida and Allshire, 2005). However, silencing of TEs may have collateral effects on the transcription of nearby genes, which can lead to preferential loss of methylated TEs from gene-rich chromosomal regions (Hollister and Gaut, 2009).

Cultivated *Brassica* species belong to the monophyletic *Brassiceae* tribe within the dicotyledon family *Brassicaceae*. Diploid *Brassica* genomes were hypothesized to have been triplicated and confirmed by many molecular marker and cytogenetic evidence (for review, Prakash et al., 2009) as well as recent whole genome sequencing of *B. rapa* (Wang et al., 2011) and *B. oleracea* (Liu et al., 2014; Parkin et al., 2014). Whole genome sequencing provides an opportunity to identify different subgenomes within the *B. rapa* genome by synthetic comparison between *B. rapa* and *A. thaliana*. Three subgenomes have been proposed, each with significant deviations from equivalent gene frequencies. The least fractionated (LF) subgenome retains 70% of the genes found in *A. thaliana*, the medium fractionated (MF1) 46%, and the most fractionated (MF2) 36% (Wang et al., 2011). Based on examining short exonic deletions in retained *Brassica* genes, Tang et al. (2012) further revealed that subgenome II (MF1) had more recent deletions than subgenome I (LF) or subgenome III (MF2), which suggested that a two-step process of genome fractionation had indeed occurred.

Here, we want to ask if biased gene loss during *B. rapa* genome evolution has been driven by, or has consequences for, epigenetic processes. In other words, if there are significantly difference in DNA methylation for genes of different copy numbers or with different subgenomes. We firstly generated a genome-scale DNA methylation profile for *B. rapa* using modified RRBS, and then compared this profile with gene transcription data from RNA-seq. We found that genes in the different subgenomes display a hierarchical level of cytosine methylation and transcription. In particular, the singleton genes have a significantly higher level of DNA methylation than genes with two or three paralogues, and are expressed at a lower level.

Materials and Methods

Samples and DNA Extraction

A semi-winter type *B. rapa* var. oleifera ($2n = 20$, AA genome, genotype 3H120) was used in this study. This inbred line had previously been used as one of the parents for new allopolyploid synthesis in order to investigate genetic and epigenetic changes following hybridization and genome doubling between different *Brassica* diploid species (Cui et al., 2012). Because all hybrid immature embryos were cultured and new plants were developed and conserved on MS medium, the parent plants used for hybridization were then also conserved by tissue culture. Briefly, the plant was firstly sub-cultured on MS medium with 1.5 mg/liter 6-benzyl aminopurine (6-BA) and 0.25 mg/liter a-naphthalenacetic acid (NAA) to generate sufficient cloned plantlets, which were then successively cultured on MS agar medium. Young leaves were collected from the young plants on MS medium and immediately frozen in liquid nitrogen. Genomic DNA was extracted from ∼100 mg tissue using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA), and DNA content was quantified by Qubit HS dsDNA kit.

Abbreviations: RRBS, reduced representation bisulfite sequencing; NAA, a-naphthalenacetic acid; 6-BA, 6-benzyl aminopurine; PE, paired-end; TSS, transcriptional start site; TRR, transcription termination region; RPKM, reads per kb per million reads; FPKM, fragments per kb per million reads; WGBS, whole-genome bisulfite sequencing; MSAP, methylation sensitive amplification polymorphism.

Bisulfite Treatment and Sequencing Library Construction

Approximately 500 ng gDNA was simultaneously double digested using *SacI* (GAGCTC) and *MseI* (TTAA) (Fermentas) in a reaction volume of 25 μl. The reaction mixture was first incubated at 37°C for 6 h, and then at 65°C for 90 min. Sac_meAD and Mse_meAD adaptors (**Figure S1**) were annealed using the program: 94°C gradually decreased to 65°C with −0.5°C every 10 s, then kept at 65°C for 10 min, 56°C for 10 min, 37°C for 10 min, and 22°C for 10 min. Restriction fragments were ligated to the Sac_meAD and Mse_meAD adaptors with unique index sequences. The ligation reaction was carried out in 50 μl at 16°C overnight with 25 pmol Sac_meAD and Mse_meAD adaptors, and 50,000 Units of T4 DNA ligase (NEB). The resulting ligates with different index sequences were mixed and concentrated using a PCR purification kit (Qiagen, Valencia, CA) and fragments between 250 and 500 bp were cut from a 2% agarose gel and purified with the Qiagen gel purification kit (Qiagen, Valencia, CA). ∼500 ng recovered products were subjected to two successive treatments with sodium bisulfite using EpiTect Bisulfite kit (Qiagen, Valencia, CA) following the manufacturer's instructions. After a final purification using the PCR purification kit, 5 μl bisulfite-converted ligates were amplified by 18 PCR cycles with the following reaction composition: 1× Taq buffer, 3.5 mM $MgCl_2$, 0.4 mM dNTPs, 1 U Taq DNA polymerase (Fermentas), and 5 pmol Illumina PCR primers (Chen et al., 2013). The enriched library was purified with Qiagen gel purification kit, and quantified by Qubit HS dsDNA kit. The library was sequenced on Hiseq 2000 platform according to the manufacturer's instructions.

Sequence Filtering and Alignment

After parsing reads into different subsets based on the index sequences, the first 75 bp of paired-end (PE) reads were retained, and the residual enzyme recognition sequences trimmed. Low-quality PE reads containing more than 5% of nucleotides with Phred quality value < 30 were filtered by the IlluQC.pl script included in NGSQCToolkit_v2.3 program suit (Patel and Jain, 2012). The remaining high-quality reads were mapped against the *B. rapa* var. *pekinensis* Chiifu-401-1 reference genome sequence (v1.2) (Wang et al., 2011) using Bismark_v0.7.4 software (Krueger and Andrews, 2011) in a non-directional manner with a maximum of 1 bp mismatch in multi-seed alignment. Only uniquely mapped reads were retained for further analyses.

Calling Methylated Loci

Overlapping sequences of paired-end reads were ignored to prevent mis-calculating the level of methylation. In order to call a methylation score for each potential CG, CHG, and CHH site, high quality cytosines ($\geq$ 20 phred quality score) within methylation loci having at least 10-fold coverage were extracted by the methCall.pl script (https://code.google.com/p/methylkit/source/browse/exec/methCall.pl?r=dd63fb95d718356e94c46ef2885d4110b385297d). Gene and TE annotations were obtained from BRAD (http://brassicadb.org/brad/). Tandem and inverted repeats were detected using Tandem Repeat Finder and Inverted Repeat Finder software packages following default parameters (Benson, 1999; Warburton et al., 2004).

SNP Detection Using a Modified ddRAD Protocol

To remove the influence of nucleotide variations during the calling of methylated loci, modified ddRAD sequencing was performed simultaneously according to the protocol published previously (Chen et al., 2013). After sequence trimming, 75 bp paired-end clean data were aligned to the *B. rapa* reference genome sequence (v1.2) using Bowtie2 software with a maximum of one mismatch (Wang et al., 2011). SNP calling was performed by Samtools software with the parameters of at least one coverage with phred quality of $\geq$ 20 (Li et al., 2009; Langmead and Salzberg, 2012). Finally methylation sites disrupted by SNPs in the 3H120 genome were excluded from further analyses.

Methylation Level Distributions Analysis

To examine the genome scale distribution of methylated and repeat sequences, we plotted the average methylation level and length of repeat sequences across each chromosome using a 200 kb sliding windows with 100 kb overlap. The length of genes and transposons were variable, hence we plotted the methylation level using a sliding windows corresponding to 10% of the length of specific genes or transposons. Promoter regions were defined as the 200 bp immediately upstream of the transcriptional start site (TSS) of each gene, and upstream and downstream of genes and transposons were defined as 1 kb 5′ and 3′.

Gene Expression Analysis

RNA-seq data from different organ and tissues of 3H120 (Zhang et al., unpublished data) and Chiifu-401 (Tong et al., 2013) were used for gene expression analysis. Genes were firstly classified into different subgenomes or groups with different copies according to published *B. rapa* reference genome (Wang et al., 2011). Then, in each group, genes were assigned into three classes of high (RPKM/FPKM $>$ 50), medium (5 $<$ RPKM/FPKM $\leq$ 50) and low (RPKM/FPKM $\leq$ 5) transcription (Tong et al., 2013) (**Table S1**).

Results

Representative DNA Methylation Profile for the *B. rapa* Genome

In order to generate a representative profile of the global DNA methylation in *B. rapa*, a modified RRBS protocol was used (**Figure S1**), yielding a total of 2.3 Gb PE100 (100 bp paired-end) sequence data. Following trimming and filtering, 1.28 Gb (corresponding to 8.56 million PE75 reads) were retained for subsequent analyses, of which 0.55 Gb (42.9%) could be successfully and uniquely aligned to the *B. rapa* reference genome using Bismark (Krueger and Andrews, 2011). These data were used to call a methylation level for each CG, CHG, and CHH site. Because SNPs between the reference *B. rapa* Chiifu-401 and the 3H120 genotype could potentially interrupt the methylation calling, we performed standard non-bisulfite sequencing of the

3H120 genome following the modified double digest Restriction-Site Associated DNA sequencing (ddRADseq) protocol (Chen et al., 2013). After sequence trimming and filtering, a total of 3.2 million PE75 (0.48 G) high-quality reads were collected. These sequences were aligned to the reference *B. rapa* Chiifu-401 genome by Bowtie2, and a total of 36,836 candidate SNPs were detected using Samtools. Of the methylated loci that had been called, 1031 CG, 736 CHG, and 1886 CHH sites were disrupted by these candidate SNPs and so were excluded. Finally, a total of 0.26 million CG, 0.27 million CHG, and 1.05 million CHH loci were recovered following alignment, which respectively accounted for 2.24, 2.16, and 1.68% of the total loci in the *B. rapa* genome. Taking those with a minimum sequencing depth of 10, the frequency of the three 5^{m}C contexts detected represented 0.64% CG, 0.59% CHG, and 0.43% of all CHH loci and were used for subsequent analyses. In order to determine whether the relative proportion of CG, CHG, and CHH sites enriched in genic and transposon regions was consistent with those in other genomes (**Figure S2A**), we performed an *in silico* mRRBS (simulation restriction enzyme digestion analysis) of the *B. rapa var. pekinesis* (Chiifu-401-1) and rice (*Oriza sativa L.SSP indica* 93-11) genome and calculated the relative proportion of each of the three methylation contexts (**Figure S2B**). We found a similar proportional representation of the frequencies at the whole genome level. This gave us confidence that the modified RRBS protocol we adopted would provide a reliable representation of DNA methylation across the *B. rapa* genome.

DNA Methylation Landscape of *B. rapa* Genome

Our sampling of DNA methylation in *B. rapa* indicated that 52.4% of CG sites were present as ^{5m}CG, with 31.8% of CHG and 8.3% of CHH (**Table S2**). At single base resolution, 92.5% of CG sites were either unmethylated or highly methylated (90–100%), whereas 71% CHH sites were either unmethylated or hypomethylated (0–10% as ^{5m}CHH per site). For the CHG context, 51.8% of sites were hypomethylated and 10.3% sites were highly methylated, with a more uniform distribution between 10 and 90% (**Figure S3**). These differences may result from the distinct genetic control under which context methylation arises and is maintained (Law and Jacobsen, 2010; He et al., 2011).

The distribution of the mean methylation level for each of the three contexts (CG, CHG, and CHH) in *B. rapa* was calculated for each chromosome, with a sliding windows of 200 kb (**Figure 1**). CG methylation was consistently higher than CHG and CHH, with CHH methylation lowest throughout the genome. We found that the average CG, CHG, and CHH methylation distributions were highly correlated, despite being maintained through distinct genetic mechanisms. In order to study the relationship between the level of DNA methylation and repeat elements, transposons, tandem and inverted repeat sequences, and the frequency of these sequence classes was plotted in sliding 200-kb windows across each chromosome (**Figure 1**). This indicated that higher levels of DNA methylation level were associated with regions enriched for repetitive sequences, and low levels of methylation distribution were dispersed in regions enriched for genes. We found a more complex distribution of methylation across chromosomes than that observed in *Arabidopsis* (Cokus et al., 2008), although there remained a positive correlation with repeat elements and a negative correlation with gene density (**Figure 1**). These results indicated that overall the distribution of DNA methylation largely reflects the relative density of transposons, retrotransposons, and other repetitive sequences.

In *Arabidopsis*, hyper DNA methylation has been associated with pericentromeric regions, primarily as a result of enrichment of diverse repetitive sequences (Cokus et al., 2008). Within the *B. rapa* genome, 18 of 21 paleocentromeric regions had been detected, including 10 extant *B. rapa* centromeres (Wang et al., 2011; Cheng et al., 2013). As expected, we found extensive DNA methylation in these repeat-rich pericentromeric regions for most chromosomes. However, a lower level of DNA methylation was associated with the pericentromeric region of chromosome A02. This reflects the reduced density of repeat sequences mapped in the vicinity of the A02 centromere. Interestingly, high levels of DNA methylation were found distributed around the eight remaining ancestral centromere regions, particularly in chromosome A02 (**Figure 1**).

Patterns of DNA Methylation in Different Components

We characterized the methylation patterns of transcribed genes and TEs in *B. rapa*, by comparing the average DNA methylation level for each context (**Figure 2A**). In genic regions, a greater than two-fold methylation was detected in introns (CG 54.4%, CHG 22.0%, and CHH 5.9%) compared to exonic regions (CG 25.1%, CHG 8.8%, and CHH 2.7%). The level of DNA methylation in exon sequences differed from intron sequences for each context, while they were similar in promoter regions (defined as 200 bp upstream of transcriptional start site, TSS) (**Figure 2A**). An approximately two-fold higher methylation was detected in TE regions (CG 88.0%, CHG 54.0%, and CHH 17.7%) compared with introns, which suggests more extensive methylation in transposons compared to genic regions.

We then characterized the methylation patterns in genic and TE regions at base-pair resolution for each context. This revealed a similar pattern of CG, CHG, and CHH methylation for these genomic components as found in *Arabidopsis*, although with a higher level of methylation in genic regions (**Figure 2B**). The levels of CG methylation in genic regions decreased from 1 kb upstream to the TSS and increased throughout the transcribed region before decreasing again up to the transcription termination region (TTR), where after it increased throughout the downstream region. A contrasting pattern was observed for non-CG methylation, with relatively low levels in the gene body compared to upstream and downstream regions, and the lowest levels of methylated detected around the TSS and TTR regions, which in *Arabidopsis* are more highly correlated with gene expression (Cokus et al., 2008), Rice (Li et al., 2012), and Soybean (Song et al., 2013).

An important function of DNA methylation in plant genomes is to modulate the silencing of transposon elements. We found that DNA methylation in TEs is indeed higher than genic regions, and also higher than upstream and downstream regions of TEs. Interestingly, the level of DNA methylation changes dramatically at the TE boundaries, with a sharp increase and

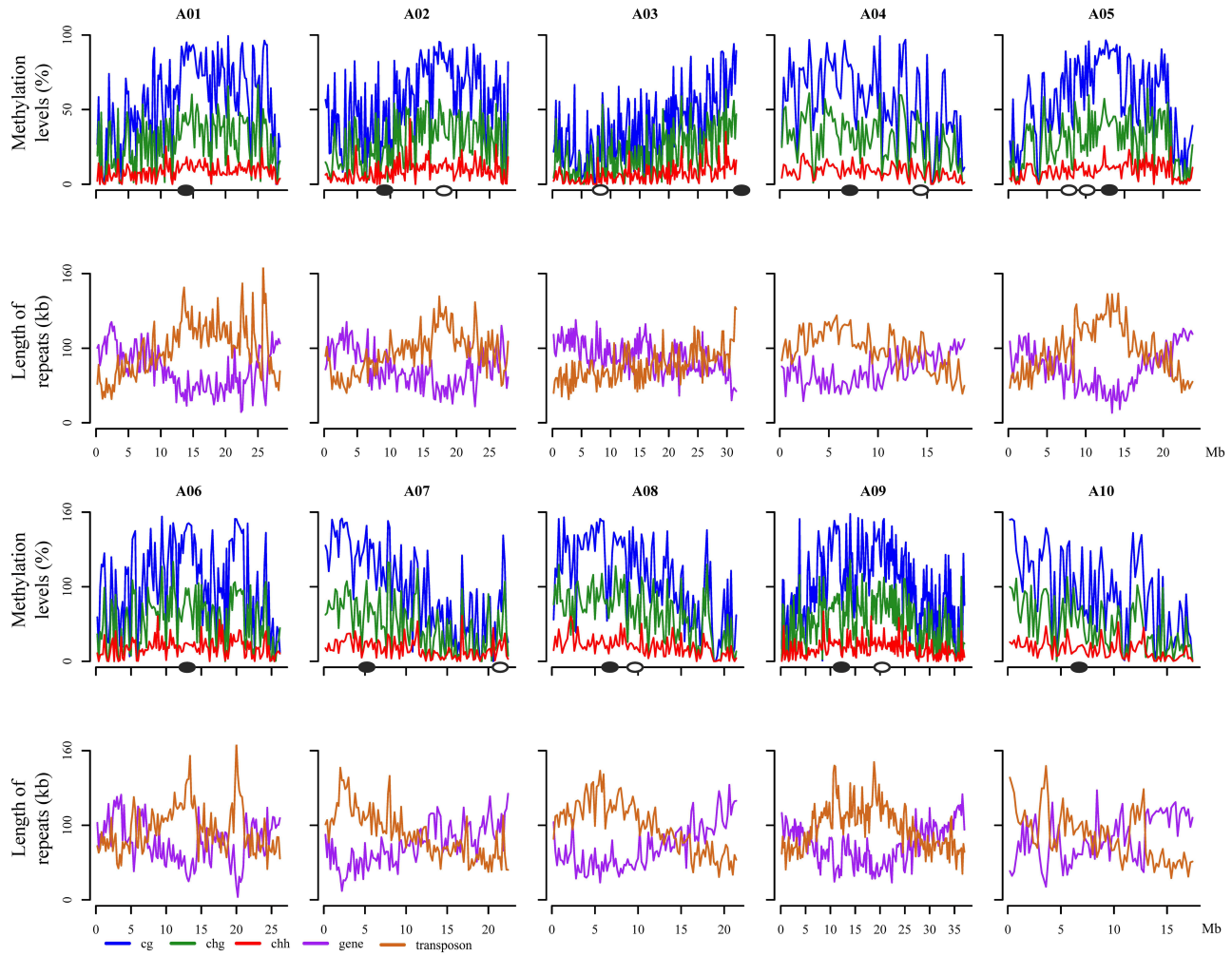


FIGURE 1 | Genome-wide distribution of methylation level, transposon and gene contents along 10 *B. rapa* chromosomes. The X axis indicates the physical distances, and Y axis indicates the average methylation level or length of repeat in each 200-kb windows. Solid oval circles stand for the position of real centromeres, blank oval circles stand for the position of ancient centromeres (Wang et al., 2011; Cheng et al., 2013). CG, CHG, and CHH (where H = A, T, or C) indicates methylated cytosine residues in different sequence contexts.

decrease around the transcription start and end sites, although DNA methylation levels are maintained relatively consistently across the transposons (**Figure 2C**).

Differential Methylation of RNA PolII Transcribed Genes

Ancestral duplication events have resulted in the mesopolyploid *B. rapa* having triplicated genomic segments, each of which have undergone different levels of gene loss (Wang et al., 2011). These three subgenomes have therefore also been defined according to their ratios of differential gene loss. The phenomenon is apparent from the retained duplication of some genes and single-copy status of others, although the mechanisms driving this are remain unclear. Three subgenomes, LF, MF1, and MF2 have been identified in the *B. rapa* whole genome sequence (Wang et al., 2011; Cheng et al., 2013). We generated profiles of DNA methylation in genic regions belonging to each subgenome. The average methylation for RNA PolII transcribed genes differed significantly between sub-genomes for each context, with CG: 29.71% (LF), 33.68% (MF1), and 31.76% (MF2); CHG: 11.23, 13.86, and 12.69%; and CHH are 3.46; 4.48, and 3.60% ($\chi^2 > 6.63$, $P < 0.01$ in each pairwise comparison). We also investigated the average methylation in each sequence context for upstream, promoter, exon, intron, and downstream components of protein coding genes (**Figure 3A**, **Figures S4A**, **S5A**). It was clear that the average methylation level of genic regions in MF1 was higher than LF and MF2, apart from downstream CHH, where LF is highest. Methylation levels in LF and MF2 appeared very similar. For the CG and CHG context, MF2 is slightly higher than LF, apart from the downstream region, although for the CHH context, LF appears slightly higher than MF2 in all components. When we classified the methylated loci according to the three subgenomes and plotted the methylation level across genic regions, we also found that MF1 was clearly more

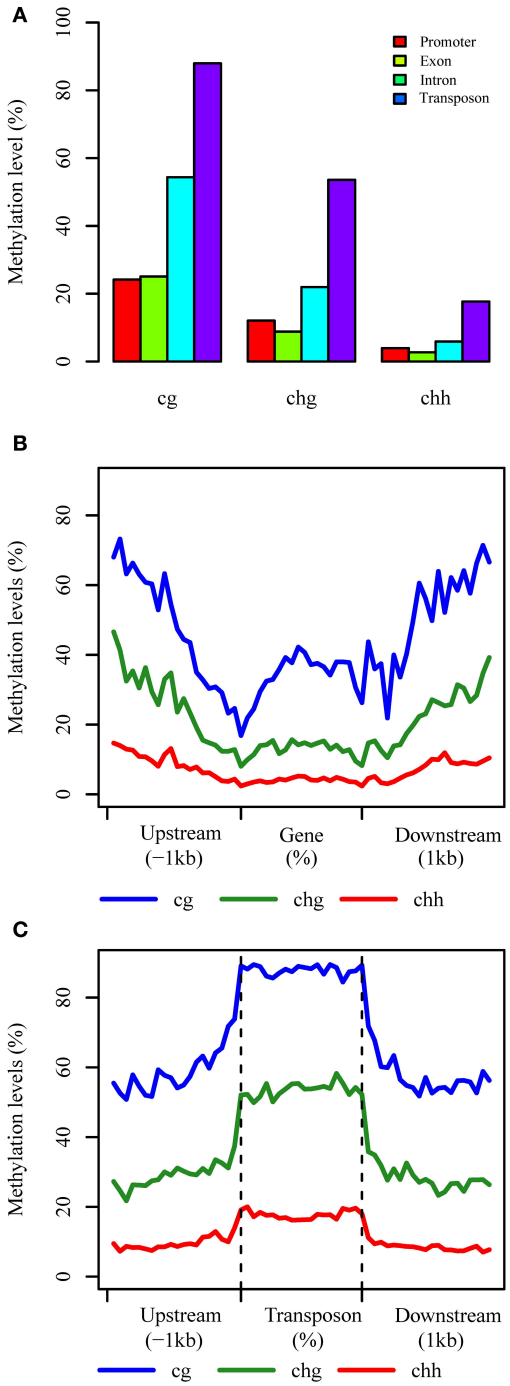


FIGURE 2 | The distributions of each context in different components. **(A)** Average methylation level in different components, and promoter regions are defined as 200 bp region before TSS; **(B,C)** Methylation patterns of genic and TE regions, respectively at base-pair resolution.

methylated than the other two subgenomes, although there were few systematic differences between LF and MF2 (**Figures 4A–C**). Overall, we conclude that genes in the MF1 subgenome are most highly methylated and in the LF subgenomes least methylated.

We also characterized average methylation levels in different component regions of single-copy, duplicated, and triplicated genes (**Figure 3B**; **Figures S4B**, **S5B**). For the CG context, we found that methylation in promoter regions of single-copy genes was 6.4-fold higher than duplicated genes and 3.9-fold higher than triplicated genes (**Figure 3B**). For duplicated and triplicated genes, a similar difference was observed, although it was inconsistent with respect to different sequence components. For example, although there is less methylation in promoter and exon regions of duplicated compared with triplicated genes, we found the converse for other components. In contrast, the pattern of methylation was relatively consistent for the CHG and CHH contexts, with only some differences appearing in promoter regions, whereas for the CHH context there were similar levels of methylation in these regions for both duplicated and triplicated genes (**Figures S4**, **S5**). In contrast to the differences in methylation observed between subgenomes, when we classified and plotted the methylated loci according to different-copy genes, very clear differences were detected across the genic regions between single-copy and duplicated genes (**Figures 4D–F**). We also observed that for the gene body, regions before the TSS and after the TTR were mostly differentially methylated, and are likely to be most responsive with respect to transcriptional repression or activation.

We further analyzed the DNA methylation of single-copy and duplicated genes in each subgenome (**Figure 5**). For single copy genes in MF1 and MF2, the level of methylation is either very similar or MF1 > MF2 apart from downstream regions. However, duplicated genes did not appear to have a systematic difference between the three subgenomes, although MF1 was clearly hypermethylated in promoter region for the CG and CHG contexts (**Figures 5A,B**). Overall, we observed a consistent difference in level of methylation level in single-copy genes between LF and the other two subgenomes, although this was not apparent in duplicated genes (**Figures S6**, **S7**). Thus, we considered that differential level of DNA methylation between different the subgenomes was primarily accounted for by the difference in single-copy genes.

Differential Transcription in Relation to DNA Methylation, Ancestral Subgenome, and Gene Copy Number

In, 3H120, for leaf tissue (**Figure 6A**), we found that more genes were included in the high and medium transcription group with fewer genes in the low group in LF than in MF1 and MF2, which indicated that genes in the LF subgenome are significantly expressed at a higher level. In contrast, for MF1 fewer genes were included in the high and medium group and more genes had low levels of transcription than in LF and MF2. Genes in MF2 showed a medium transcription level between MF1 and LF. Thus, the, average transcription level in three subgenome fitted the LF > MF2 > MF1 relationships. We also characterized transcription in leaf tissue for different gene copies (**Figure 6B**) and found that only 7.3% of single-copy genes were highly expressed, compare to 14.4% three-copy genes and 10.4% two-copy genes. In the medium expression group, the percentage of two-copy genes is highest (37.0%), while lowest (30.7%) for single-copy genes.

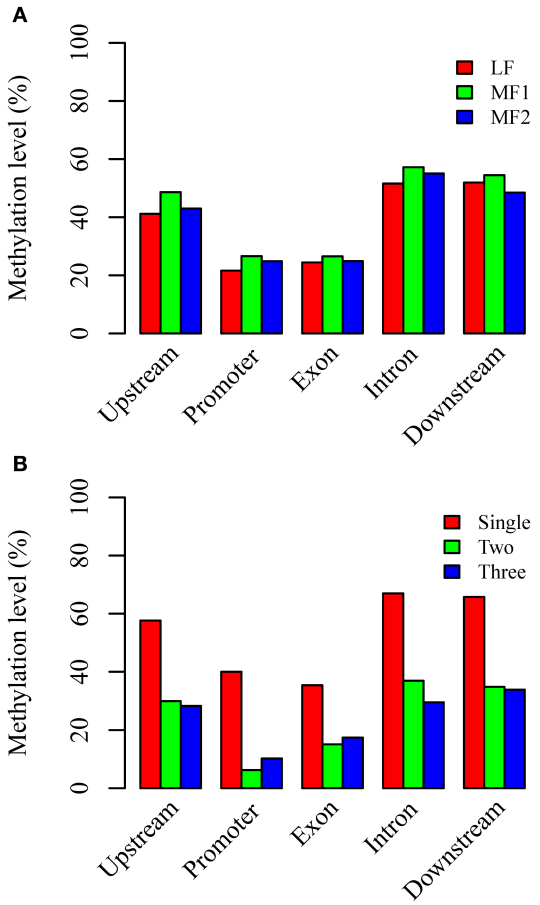


FIGURE 3 | Average CG methylation levels in different components of genic regions (A) between three subgenomes and (B) between different copy genes. Whereas, promoter regions are defined as 200 bp regions before TSS site, upstream or downstream regions are defined as 1000 bp before TSS or after TTR sites.

Compared to two- or three-copy genes, the majority of single-copy genes had a relatively lower level of transcription.

This pattern appeared to be consistent with results for silique tissue from *B. rapa* 3H120 and in different tissues from Chiifu-401-42 (data from Tong et al., 2013) (**Table S3**). We therefore deduce that the level of transcription for duplicated genes is significantly dominant compared to single-copy genes, and that this is consistent with the significantly higher DNA methylation of single-copy genes compared with duplicated genes. In conclusion, we find strong evidence that the relationship between DNA methylation and level of transcription is more dependent upon the copy number of genes than it is between different subgenomes of *B. rapa*.

Discussion

Cytosine methylation is an epigenetic modification of DNA that is also associated with histone modification and nucleosome positioning. In higher plants, methylation modification of cytosine (5^mC) is present in CG, CHG, and asymmetric CHH sequence contexts (where H is A, C, or T) (Henderson and Jacobsen, 2007). Whole-genome bisulfite sequencing (WGBS), including MethylC-seq (Lister et al., 2009) and BS-seq (Cokus et al., 2008; Laurent et al., 2010), is the most comprehensive method for genome wide methylation analysis giving single-cytosine methylation resolution and direct estimates of the proportion of molecules methylated. However, the method requires deep re-sequencing of the entire genome, which is still expensive for complex crop genomes.

The common RRBS protocol tends to enrich CG-rich sequences in the genome due to the usage of restriction enzyme MspI. For mammalian genomes, this enables the majority of CG islands, promoters or other relevant genomic regions in to be captured with limited sequence data (Gu et al., 2011). However, plants have a different pattern of CG distribution, "mosaic" methylation across genome and lack of high unmethylated CG islands (Feng et al., 2010). Moreover, this method is also limited by the uneven distribution of captured regions across chromosomes and inability to represent all sequence components, rendering it unsuitable for profiling genome-wide DNA methylation. In our previous study, SacI/MseI RE combination was used to construct a modified ddRAD library for SNP calling in a doubled haploid (DH) population, and the targeted fragments were found to be evenly and randomly distributed across the *Brassica* genome (Chen et al., 2013). Hence a new double enzyme digested RRBS method was used here to interpret the global DNA methylation at single-base resolution in *B. rapa*. The results show that the percentage of CG, CHG, and CHH loci located in genic regions was consistent between enriched targeted regions using modified RRBS protocol and whole-genome methylation loci. Coupled with an *in silico* double digestion analysis of the rice genome, we were able to confirm the applicability of this modified RRBS approach. Due to the advantages of cost effectiveness and simplicity, modified RRBS is well-suited for DNA methylation profiling of large natural populations or for construction DNA methylation genetic maps (Long et al., 2011).

The methylation ratios observed for each context in *B. rapa* were much higher compared to those reported for the DNA methylome of *Arabidopsis* (Cokus et al., 2008). This may have resulted from the higher level of repeat sequences, especially with for the additional transposable elements in *B. rapa* genome compared with *Arabidopsis* (Wang et al., 2011). In comparison with recently released *B. oleracea* (C genome) and *B. napus* (AC) methylomes, our *B. rapa* analysis has indicated a similar percentage of mCG but higher levels of mCHG and mCHH. It is perhaps surprising that the methylation level of the three contexts are similar or higher in this *B. rapa* compared with *B. oleracea* which contains significantly more repeat sequences, specific recent expansion of the *Bot1* CACTA transposon family (Alix et al., 2008) and a reported higher level of methylation (Liu et al., 2014). In *B. napus*, the C_n subgenome was also found to have a higher methylation level than that of the A_n subgenome (Chalhoub et al., 2014). Our results suggest a slight increase on this for the *B. rapa* methylome. This may have resulted from the different methods used, with only 2% of the complete set

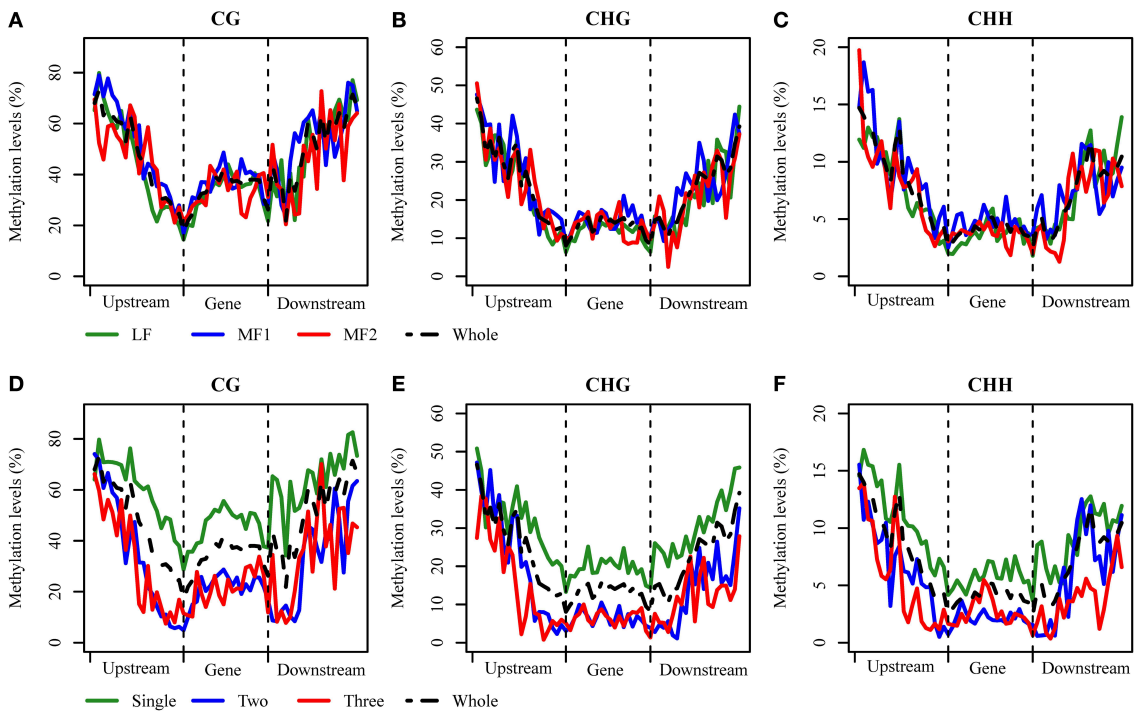


FIGURE 4 | DNA methylation distributed across genic regions at base-pair resolution. (A–C) between three subgenomes for CG, CHG, and CHH context and **(D–F)** between different copy genes for CG, CHG, and CHH context.

of genome loci recovered in this analysis. We also note that our DNA was isolated from tissue cultured plants. Tissue culture has been found to induce DNA methylation changes in different plant species (Kaeppler and Phillips, 1993; Hang et al., 2009; Linacero et al., 2011; Gonzalez et al., 2013; Stroud et al., 2013; Stelpflug et al., 2014), with the degree and the direction of methylation changes varying with different tissues or cell-types and culture methods. In *Arabidopsis* suspension culture, it was found that a prevalence of DNA methylation increases in genic regions as opposed to losses (Bednarek et al., 2007; Tanurdzic et al., 2008). In recent whole genome level surveys of maize and rice, it was found that losses of DNA methylation following tissue culture are more common than gains of DNA methylation. Meanwhile, it was also found that the bulk of the methylome were not affected, although a subset of genomic regions exhibit altered DNA methylation levels (Stroud et al., 2013; Stelpflug et al., 2014). Although we do not know the true effect on cultured *B. rapa* here, the culture methods used here were similar to those of mazie and rice (Stroud et al., 2013; Stelpflug et al., 2014). Moreover, an earlier methylation sensitive amplification polymorphism (MSAP) survey of the *B. napus* genome (Long et al., 2011) found that a very high proportion of parental methylation alleles were conserved intact in segregating lines maintained through five meiosis following initial tissue culture of the F1 line.

A "two-step theory" for paleohexaploid *B. rapa* formation is sufficient to explain why MF1 and MF2 are more fractionated than the LF subgenome (Wang et al., 2011; Cheng et al., 2013). It has been proposed that MF1 and MF2 are of similar age and first came together to form a tetraploid, with subsequent inclusion of LF to form the ancestral hexaploid. MF1 and MF2 may therefore have resided in the same nucleus for a longer period of time than LF, which is then relatively less fractionated than the first two. We were interested to establish whether the relative methylation level of these different subgenomes may provide some insights into an epigenetic basis for complex genome evolution (Schnable and Freeling, 2011; Diez et al., 2014; Woodhouse et al., 2014). As anticipated, we found that genes in LF had the lowest level of methylation, at least for CG and CHG contexts, corresponding to the highest level of gene transcription. These results are consistent with those in *B. oleracea*, in which lower methylation levels were found in the least fractionated genome (Parkin et al., 2014), although the levels for MF1 and MF2 were reversed with respect to *B. rapa*. We found that for *B. rapa* the levels of methylation were inversely related to gene expression for each subgenome (DNA methylation: MF1 > MF2 > LF; Gene expression: LF > MF2 > MF1), with a bias to fractionation in MF2 compared with MF1 that was not consistent with the pattern of epigenetic marks. It is most likely that MF1 and MF2 came together and contributed to an early tetraploid karyotype (at least 5–9 MYA) (Wang et al., 2011), and over a long period of time, MF2 emerged as the subdominant genome, carrying a higher load of DNA methylation and associated lower level of gene transcription, which resulted in greater gene loss compared with MF1. However, following the incorporation of LF, the original methylation status may have been modified by epigenetic reprogramming in the early stages of the new

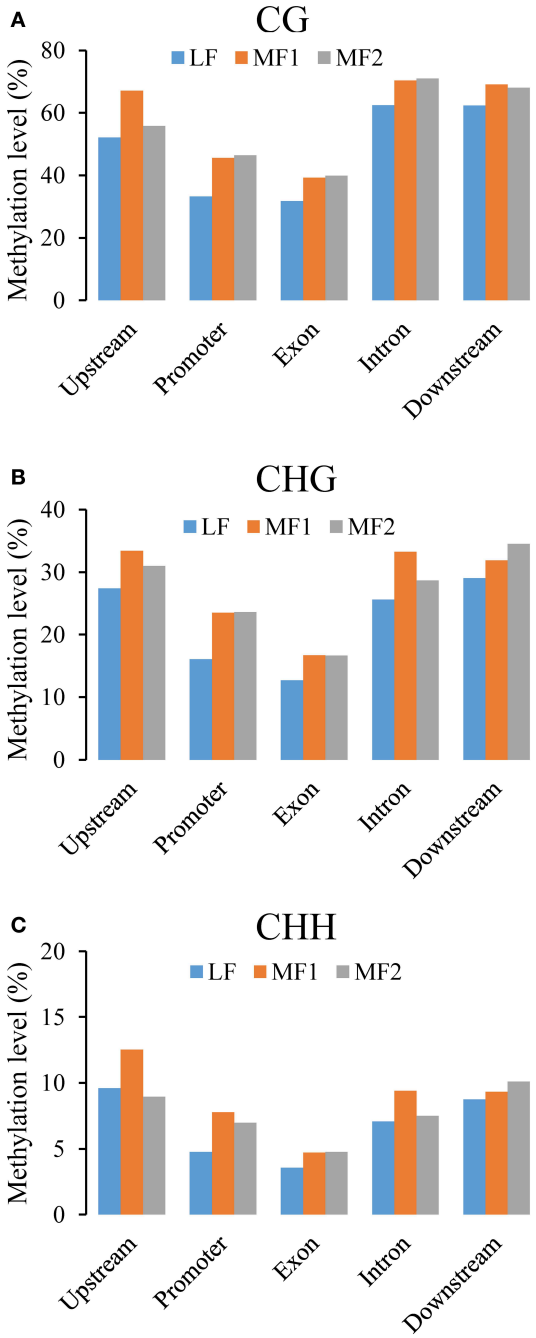


FIGURE 5 | Average methylation level in single-copy genes between three subgenomes for (A) CG, (B) CHG, and (C) CHH context. Whereas, promoter regions are defined as 200 bp regions before TSS site, upstream or downstream regions are defined as 1000 bp before TSS or after TTR sites.

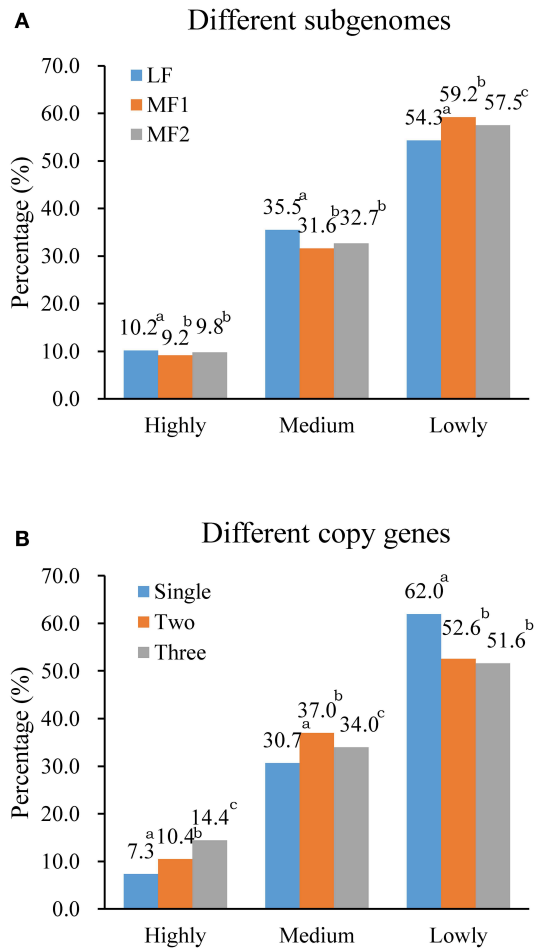


FIGURE 6 | Differential gene expression (A) between three subgenomes and (B) between different copy genes. Expressed genes were collected into lowly (≤ 5), medium (5 < RPKM ≤ 50) and highly (> 50) groups according to RPKM value. Superscript a, b, c indicated the significance of gene expression level with different copies or within different subgenomes.

polyploid formatting (Lukens et al., 2006; Gaeta et al., 2007; Szadkowski et al., 2010; Cui et al., 2013). This provides a realistic explanation for the resulting MF1 > MF2 > LF hierarchy of methylation. Subsequently, a new cycle of gene loss is likely to have arisen, based on the revised pattern gene expression. This explains the more recent loss of sequence from MF1 (Tang et al., 2012). However, we are of course aware that there is many other contributing to variation in DNA methylation, including interdependent relationships between genic methylation and transcription (Zilberman et al., 2007). Additional data from a wider range of representative sub-taxa would help confirm these hypotheses based upon our initial survey.

In contrast to the complex pattern of methylation in different subgenomes, we found that methylation level in single-copy genes was universally higher than duplicated genes, with a correspondingly lower level of expression. These results are consistent with a lower level of transcription contributing to genes being more readily deleted due to lower fitness, so that they escape purifying selection (Diez et al., 2014). This further substantiates the hypothesis that epigenetic processes are a viable alternative mechanism for duplicated gene loss. Meanwhile, we find that the methylation level of single copy genes alone is good indicators of the methylation status of any given subgenome. Consistent with this, the methylation of the three ancestral

sub-genomes of *B. oleracea* did not appear to be reflected at the level of the retained triplicate genes (Liu et al., 2014). The methylation and gene expression of retained single copy genes thus provide more credible residue markers of the intrinsic status of ancestral genomes, with biased gene loss during the formation and evolution of polyploids.

Conclusion

We generated a representative whole genome methylation profile for the first time in *B. rapa* by using a modified RRBS method. We found that methylation level generally reflected the dominance of gene loss and gene expression between different ancestral subgenomes. The results here provide more evidence for the involvement of epigenetic mechanisms in polyploid genome evolution, as well as alternative mechanism for determining the fate of duplicated genes.

Author Contributions

The study was conceived by XC, XG, and KL. CT prepared the plant materials. XC and JW performed the experiments. XC and XG contributed to data analysis, bioinformatics analysis, and manuscript preparation. GK participated in writing the manuscript. All authors contributed to revising the manuscript. All authors had read and approved the final manuscript.

Acknowledgments

We thank Qinghua Zhang, Huazhong Agricultural University, for help in production of Illumina sequence data. This work was supported by the Grants from Natural Science Foundation of China (31171583) to XG and Fundamental Research Funds for the Central Universities (2014QC025) to JW. GK is supported by the Hubei province Chutian Scholar programme.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00836

Table S1 | Characterization of gene expression with different copies or within different subgenomes.

Table S2 | Summary of methylated loci and annotation.

Table S3 | The relative proportion of genes transcribed at high, medium and low levels in different tissues of *B. rapa*.

Figure S1 | The construction of a modified reduced representation bisulfite sequencing library.

Figure S2 | The percentage of the three methylation contexts between whole genome and loci enriched using (*in silico*) mRRBS in (A) *B. rapa* and (B) rice (*Oryza sativa*).

Figure S3 | The ratio of genome-wide methylation levels for CG, CHG and CHH contexts.

Figure S4 | Mean CHG methylation levels in different components of genic regions (A) within the three subgenomes and (B) between genes of different copy number.

Figure S5 | Mean CHH methylation levels in different components of genic regions (A) within the three subgenomes and (B) between genes of different copy number.

Figure S6 | Mean methylation level of two-copy genes in three subgenomes of *B. rapa*.

Figure S7 | Mean methylation level of three-copy genes in three subgenomes of *B. rapa*.

Availability of Supporting Data

All the sequencing data used in this research have been submitted to public database NCBI under PRJNA281682.

References

Alix, K., Joets, J., Ryder, C. D., Moore, J., Barker, G. C., Bailey, J. P., et al. (2008). The CACTA transposon *Bot1* played a major role in *Brassica* genome divergence and gene proliferation. *Plant J.* 56, 1030–1044. doi: 10.1111/j.1365-313X.2008.03660.x

Almeida, R., and Allshire, R. C. (2005). RNA silencing and genome regulation. *Trends Cell Biol.* 15, 251–258. doi: 10.1016/j.tcb.2005.03.006

Bednarek, P. T., Orłowska, R., Koebner, R. M., and Zimny, J. (2007). Quantification of the tissue-culture induced variation in barley (*Hordeum vulgare L.*). *BMC Plant Biol.* 7:10. doi: 10.1186/1471-2229-7-10

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27, 573–580. doi: 10.1093/nar/27.2.573

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21. doi: 10.1101/gad.947102

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chan, S. W., Henderson, I. R., and Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* 6, 351–360. doi: 10.1038/nrg1601

Chen, X., Li, X., Zhang, B., Xu, J., Wu, Z., Wang, B., et al. (2013). Detection and genotyping of restriction fragment associated polymorphisms in polyploid crops with a pseudo-reference sequence: a case study in allotetraploid *Brassica napus*. *BMC Genomics* 14:346. doi: 10.1186/1471-2164-14-346

Cheng, F., Mandáková, T., Wu, J., Xie, Q., Lysak, M. A., and Wang, X. (2013). Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *Plant Cell* 25, 1541–1554. doi: 10.1105/tpc.113.110486

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., et al. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219. doi: 10.1038/nature06745

Cui, C., Ge, X., Gautam, M., Kang, L., and Li, Z. (2012). Cytoplasmic and genomic effects on meiotic pairing in *Brassica* hybrids and allotetraploids from pair crosses of three cultivated diploids. *Genetics* 191, 725–738. doi: 10.1534/genetics.112.140780

Cui, C., Ge, X., Zhou, Y., Li, M., and Li, Z. (2013). Cytoplasmic and genomic effects on non-meiosis-driven genetic changes in *Brassica* hybrids and allotetraploids from pairwise crosses of three cultivated diploids. *PLoS ONE* 8:e65078. doi: 10.1371/journal.pone.0065078

Diez, C. M., Roessler, K., and Gaut, B. S. (2014). Epigenetics and plant genome evolution. *Curr. Opin. Plant Biol.* 18, 1–8. doi: 10.1016/j.pbi.2013.11.017

Feng, S., Cokus, S. J., Zhang, X., Chen, P. Y., Bostick, M., Goll, M. G., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8689–8694. doi: 10.1073/pnas.1002720107

Gaeta, R. T., Pires, J. C., Iniguez-Luy, F., Leon, E., and Osborn, T. C. (2007). Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. *Plant Cell* 19, 3403–3417. doi: 10.1105/tpc.107.054346

Gonzalez, A., Saiz, A., Acedo, A., Ruiz, M., and Polanco, C. (2013). Analysis of genomic DNA methylation patterns in regenerated and control plants of rye (*Secale cereale* L.). *Plant Growth Regul.* 70, 227–236. doi: 10.1007/s10725-013-9794-7

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* 6, 468–481. doi: 10.1038/nprot.2010.190

Hang, M., Xu, C., Yan, H., Zhao, N., Von Wettstein, D., and Liu, B. (2009). Limited tissue culture−induced mutations and linked epigenetic modifications in F1 hybrids of sorghum pure lines are accompanied by increased transcription of DNA methyltransferases and 5-methylcytosine glycosylases. *Plant J.* 57, 666–679. doi: 10.1111/j.1365-313X.2008.03719.x

He, X. J., Chen, T., and Zhu, J. K. (2011). Regulation and function of DNA methylation in plants and animals. *Cell Res.* 21, 442–465. doi: 10.1038/cr.2011.23

Henderson, I. R., and Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418–424. doi: 10.1038/nature05917

Hollister, J. D., and Gaut, B. S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. *Genome Res.* 19, 1419–1428. doi: 10.1101/gr.091678.109

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature* 473, 97–100. doi: 10.1038/nature09916

Kaeppler, S. M., and Phillips, R. L. (1993). Tissue culture-induced DNA methylation variation in maize. *Proc. Natl. Acad. Sci. U.S.A.* 90, 8773–8776. doi: 10.1073/pnas.90.19.8773

Krueger, F., and Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. doi: 10.1093/bioinformatics/btr167

Langham, R. J., Walsh, J., Dunn, M., Ko, C., Goff, S. A., and Freeling, M. (2004). Genomic duplication, fractionation and the origin of regulatory novelty. *Genetics* 166, 935–945. doi: 10.1534/genetics.166.2.935

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Meth.* 9, 357–359. doi: 10.1038/nmeth.1923

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C. T., et al. (2010). Dynamic changes in the human methylome during differentiation. *Genome Res.* 20, 320–331. doi: 10.1101/gr.101907.109

Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, X., Zhu, J., Hu, F., Ge, S., Ye, M., Xiang, H., et al. (2012). Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* 13:300. doi: 10.1186/1471-2164-13-300

Linacero, R., Rueda, J., Esquivel, E., Bellido, A., Domingo, A., and Vázquez, A. M. (2011). Genetic and epigenetic relationship in rye, *Secale cereale* L., somaclonal variation within somatic embryo-derived plants. *In Vitro Cell Dev. Plant* 47, 618–628. doi: 10.1007/s11627-011-9407-y

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., et al. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 133, 523–536. doi: 10.1016/j.cell.2008.03.029

Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322. doi: 10.1038/nature08514

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930. doi: 10.1038/ncomms4930

Long, Y., Xia, W., Li, R., Wang, J., Shao, M., Feng, J., et al. (2011). Epigenetic QTL mapping in *Brassica napus*. *Genetics* 189, 1093–1102. doi: 10.1534/genetics.111.131615

Lukens, L. N., Pires, J. C., Leon, E., Vogelzang, R., Oslach, L., and Osborn, T. (2006). Patterns of sequence loss and cytosine methylation within a population of newly resynthesized *Brassica napus* allopolyploids. *Plant Physiol.* 140, 336–348. doi: 10.1104/pp.105.066308

Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., et al. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol.* 15:R77. doi: 10.1186/gb-2014-15-6-r77

Patel, R. K., and Jain, M. (2012). NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS ONE* 7:e30619. doi: 10.1371/journal.pone.0030619

Prakash, S., Bhat, S. R., Quiros, C. F., Kirti, P. B., and Chopra, V. L. (2009). *Brassica* and its close allies: cytogenetics and evolution. *Plant Breed. Rev.* 31, 21. doi: 10.1002/9780470593783.ch2

Proost, S., Pattyn, P., Gerats, T., and Van de Peer, Y. (2011). Journey through the past: 150 million years of plant genome evolution. *Plant J.* 66, 58–65. doi: 10.1111/j.1365-313X.2011.04521.x

Sankoff, D., Zheng, C., and Zhu, Q. (2010). The collapse of gene complement following whole genome duplication. *BMC Genomics* 11:313. doi: 10.1186/1471-2164-11-313

Schnable, J. C., and Freeling, M. (2011). Genes identified by visible mutant phenotypes show increased bias toward one of two subgenomes of maize. *PLoS ONE* 6:e17855. doi: 10.1371/journal.pone.0017855

Schnable, J. C., Springer, N. M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4069–4074. doi: 10.1073/pnas.1101368108

Soltis, D. E., and Soltis, P. S. (1999). Polyploidy: recurrent formation and genome evolution. *Trends Ecol. Evol.* 14, 348–352.

Song, Q. X., Lu, X., Li, Q. T., Chen, H., Hu, X. Y., Ma, B., et al. (2013). Genome-wide analysis of DNA methylation in soybean. *Mol. Plant*. 6, 1961–1974. doi: 10.1093/mp/sst123

Stelpflug, S. C., Eichten, S. R., Hermanson, P. J., Springer, N. M., and Kaeppler, S. M. (2014). Consistent and heritable alterations of DNA methylation are induced by tissue culture in maize. *Genetics* 198, 209–218. doi: 10.1534/genetics.114.165480

Stroud, H., Ding, B., Simon, S. A., Feng, S., Bellizzi, M., Pellegrini, M., et al. (2013). Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife* 2:e00354. doi: 10.7554/elife.00354

Szadkowski, E., Eber, F., Huteau, V., Lodé, M., Huneau, C., Belcram, H., et al. (2010). The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytol.* 186, 102–112. doi: 10.1111/j.1469-8137.2010.03182.x

Tang, H., Woodhouse, M. R., Cheng, F., Schnable, J. C., Pedersen, B. S., Conant, G., et al. (2012). Altered patterns of fractionation and exon deletions in *Brassica rapa* support a two-step model of paleohexaploidy. *Genetics* 190, 1563–1574. doi: 10.1534/genetics.111.137349

Tanurdzic, M., Vaughn, M. W., Jiang, H., Lee, T. J., Slotkin, R. K., Sosinski, B., et al. (2008). Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol.* 6:e302. doi: 10.1371/journal.pbio.0060302

Thomas, B. C., Pedersen, B., and Freeling, M. (2006). Following tetraploidy in an *Arabidopsis* ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Res.* 16, 934–946. doi: 10.1101/gr.4708406

Tong, C., Wang, X., Yu, J., Wu, J., Li, W., Huang, J., et al. (2013). Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in *Brassica rapa*. *BMC Genomics* 14:689. doi: 10.1186/1471-2164-14-689

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43:1035–1039. doi: 10.1038/ng.919

Warburton, P. E., Giordano, J., Cheung, F., Gelfand, Y., and Benson, G. (2004). Inverted repeat structure of the human genome: the X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes. *Genome Res.* 14, 1861–1869 doi: 10.1101/gr.2542904

Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nat. Rev. Genet.* 2, 333–341. doi: 10.1038/35072009

Woodhouse, M. R., Cheng, F., Pires, J. C., Lisch, D., Freeling, M., and Wang, X. (2014). Origin, inheritance, and gene regulatory consequences of genome dominance in polyploids. *Proc. Natl. Acad. Sci. U.S.A.* 111, 5283–5288. doi: 10.1073/pnas.1402475111

Woodhouse, M. R., Schnable, J. C., Pedersen, B. S., Lyons, E., Lisch, D., Subramaniam, S., et al. (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homeologs. *PLoS Biol.* 8:e1000409. doi: 10.1371/journal.pbio.1000409

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* 39, 61–69. doi: 10.1038/ng1929

Conflict of Interest Statement: The reviewer Maoteng Li declares that, despite having previously collaborated with the co-author Xianhong Ge, the review process was conducted objectively. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Chen, Ge, Wang, Tan, King and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 March 2016
doi: 10.3389/fpls.2016.00311

Comparative Leaves Transcriptome Analysis Emphasizing on Accumulation of Anthocyanins in *Brassica*: Molecular Regulation and Potential Interaction with Photosynthesis

Muhammad A. Mushtaq, Qi Pan, Daozong Chen, Qinghua Zhang, Xianhong Ge* and Zaiyun Li

National Key Laboratory of Crop Genetic Improvement, National Center of Oil Crop Improvement (Wuhan), College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China

OPEN ACCESS

Edited by:
Om Parkash Dhankher, University of Massachusetts Amherst, USA

Reviewed by:
Maoteng Li, Huazhong University of Science and Technology, China
Zhongyun Piao, Shenyang Agricultural University, China

****Correspondence:***
Xianhong Ge
gexianhong@mail.hzau.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *04 November 2015*
Accepted: *29 February 2016*
Published: *18 March 2016*

Citation:
Mushtaq MA, Pan Q, Chen D, Zhang Q, Ge X and Li Z (2016) Comparative Leaves Transcriptome Analysis Emphasizing on Accumulation of Anthocyanins in Brassica: Molecular Regulation and Potential Interaction with Photosynthesis. Front. Plant Sci. 7:311. doi: 10.3389/fpls.2016.00311

The purple leaf pigmentation mainly associated with anthocyanins accumulation is common in *Brassica* but the mechanisms of its production and its potential physiological functions are poorly understood. Here, we performed the phenotypic, cytological, physiological, and comparative leaves transcriptome analyses of 11 different varieties belonging to five *Brassica* species with purple or green leaves. We observed that the anthocyanin was accumulated in most of vegetative tissues in all species and also in reproduction organs of *B. carinata*. Anthocyanin accumulated in different part of purple leaves including adaxial and abaxial epidermal cells as well as palisade and spongy mesophyll cells. Leave transcriptome analysis showed that almost all late biosynthetic genes (LBGs) of anthocyanin, especially *Dihydroflavonol 4-Reductase* (*DFR*), *Anthocyanidin Synthase* (*ANS*) and *Transparent Testa 19* (*TT19*), were highly up-regulated in all purple leaves. However, only one of transcript factors in anthocyanin biosynthesis pathway, *Transparent Testa* 8 (*TT8*), was up regulated along with those genes in all purple leaves, indicating its pivotal role for anthocyanin production in *Brassica*. Interestingly, with the up-regulation of genes for anthocyanin synthesis, Cytosolic 6-phosphogluconolactonase (*PLG5*) which involved in the oxidative pentose-phosphate pathway was up-regulated in all purple leaves and three genes *FTSH PROTEASE 8* (*FTS8*), *GLYCOLATE OXIDASE 1* (*GOX1*), and *GLUTAMINE SYNTHETASE 1;4* (*GLN1;4*) related to degradation of photo-damaged proteins in photosystem II and light respiration were down-regulated. These results highlighted the potential physiological functions of anthocyanin accumulation related to photosynthesis which might be of great worth in future.

Keywords: *Brassica*, pigmentation, transcriptome, anthocyanins, photosynthesis

INTRODUCTION

Anthocyanins are water soluble pigment existing in many plants, algae, and bacteria. These are responsible for various color formation in leaves, flowers, stems, roots, and many other plant organs which usually attract pollinators and dispersers. Anthocyanins may also play roles in protecting chloroplast from the photo-oxidative and photo-inhibitory damage by scavenging free radicals and reactive oxygen species (ROS; Hughes et al., 2005; Hatier and Gould, 2008). Plants under stress conditions or infection by pathogens could also induce anthocyanins formation (Chalker-Scott, 1999; Lea et al., 2007; Kerio, 2011). Therefore, these pigments are highly essential for plant survival. Moreover, recent studies in tomato have indicated that accumulation of the anthocyanins on skin could double its shelf-life by delaying over-ripening and reducing the susceptibility to gray mold (Bassolino et al., 2013; Zhang et al., 2013, 2015b). A number of studies have suggested that the food with rich anthocyanins could benefit human health by its high antioxidant activity against cardiovascular disease, certain cancer, and some other chronic diseases (Hou, 2003; Butelli et al., 2008; Martin et al., 2011; Lila, 2004).

Anthocyanins are formed by phenylpropanoid metabolism from phenylalanine by series genes including early biosynthetic genes (EBGs) and late biosynthetic genes (LBGs) (**Figure 1**). Simply, three molecules Malonyl CoA and one of ρ-coumaroyl CoA are firstly condensed by chalcone synthase. The product 4, 2′4′6′-tetrahydrocychalcone are further catalyzed successively by four enzymes (Chalcone Isomerase [CHI], Flavanone 3-Hydroxylase [F3H], Dihydroflavonol 4-Reductase [DFR] and Anthocyanidin Synthase [ANS/LDOX]; Harborne and Grayer, 1994; Bohm, 1998; Dao et al., 2011). In addition to pelargonidin, two other anthocyanidin, cyanidin, and delphinidin, are formed in most plants by further hydroxylate the B-ring of dihydrokaempferol by Flavonoid 3′Hydroxylase [F3′H] and Flavonoid 3′5′ Hydroxylase [F3′5′H], respectively. Depending on variable cell environment, especially the vacuolar pH value, pelargonidin being orange to red, cyanidin being red to red-purple, and delphinidin being red-purple to blue. However, these pigments accumulate exclusively as glycosylated forms (anthocyanins) for the anthocyanidin structure are unstable. Till date, more than 600 natural anthocyanins have been identified which are derived from these core anthocyanins by methylated on its 3′ or 5′ or both hydroxyl group and side chain decorations, such as glycosylation and acylation (Glover and Martin, 2012).

The regulation of anthocyanin biosynthesis at transcription level is mainly operated by a series of transcript factors, especially for those R2R3-MYB genes. While early biosynthesis genes are active by co-activator independent and functionally redundant R2R3-MYB genes (e.g., MYB11, MYB12, and MYB111 in *Arabidopsis*), late biosynthesis genes are active by a highly conserved transcriptional activation complex MYB–bHLH–WDR (MBW) in angiosperms and likely also in gymnosperm (**Figure 1**; Hichri et al., 2011; Petroni and Tonelli, 2011; Xu et al., 2013, 2015). The complex consists of MYB proteins, basic Helix-Loop-Helix (bHLH) proteins, and a WD repeat protein. In *Arabidopsis*, R2R3-MYB genes *PAP1*, *PAP2*, *MYB113,* and *MYB114* (Borevitz et al., 2000; Gonzalez et al., 2008), *bHLH* gene family *TT8*, *GL3*, *EGL3* (Nesi et al., 2000; Payne et al., 2000; Zhang et al., 2003), and *WD40* family gene *TTG1* (Walker et al., 1999) are recognized as the key genes encoding respective components of MBW complex. In addition, various MYB TFs (i.e., MYBL2 and MYB4), SPL9 and LBD family genes are also reported as negative regulators of the anthocyanins accumulation (Jin et al., 2000; Dubos et al., 2008; Matsui et al., 2008; Rubin et al., 2009; Gou et al., 2011) (**Figure 1**).

Cultivated *Brassica* species belong to the monophyletic *Brassiceae* tribe within the dicotyledon family *Brassicaceae*, including three diploids *B. rapa* (AA), *B, oleracea* (CC), *B. nigra* (BB), and three allotetraploids *B. napus* (AACC), *B. juncea* (AABB), *B. carinara* (BBCC). Three allotetraploids are formed and evolved from the hybrids between any two of diploids (Nagaharu, 1935). Most of these species are important oilseed crops and vegetables worldwide. Although all *Brassica* species contain yellow flowers due to the accumulation of carotenoid (Zhang et al., 2015a), red and purple pigments are deposited at various part of *Brassica* plants especially at leaves, stems, and pods. Although a number of studies have been carried out in different *Brassica* species, the mechanisms behind such color formation are still poorly understood. In present study, we systemically investigated the pigments formation in five *Brassica* species and performed comparative transcriptome analysis between purple and green leaves. Our main objective was to examine potential key genes response for leaf pigmentation as well as the physiological roles of anthyocyanin accumulation in *Brassica* plants development.

Abbreviations: PAL, Phenylalanine ammonia-lyase; C4H, Cinnamate-4-hydroxylase; 4CL, 4-coumarate:CoA ligase; CHS, Chalcone synthase; CHI, Chalcone Isomerase; F3H, Flavanone 3-Hydroxylase; F3′H, Flavonoid 3′Hydroxylase; F3′5′H, Flavonoid 3′5′ Hydroxylase; FLS, Flavonol synthase; DFR, Dihydroflavonol 4-Reductase; ANS, Anthocyanidin Synthase; TT19, Transparent Testa 19/Glutathione Transferase; TT8, Transparent Testa 8; TT2, Transparent Testa 2; UFGT, UDP glucose-flavonoid 3-O-glucosyltransferase; TFs, Transcription Factors; MBW, MYB–bHLH–WDR; bHLH, basic Helix-Loop-Helix; WDR, WD repeat; PAP1, Production of Anthocyanin Pigment 1; PAP2, Production of Anthocyanin Pigment 2; EGL3, Enhancer of Glabrous 3; TTG1, Transparent Testa Glabrous 1; GL3, Glabrous 3; CPC, CAPRICE; MYBL2, MYB-Like 2; LBD, LOB Domain-containing Protein; SPL9, SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 9; PGL5, Cytosolic 6-Phosphogluconolactonase; OPPP, Oxidative Pentose-Phosphate Pathway; PSII, Photosystem II; GOX1,Glycolate Oxidase; GLN1, Glutamine Synthetase; COR27, Cold Regulated gene 27; ROS, Reactive Oxygen Species; LSD, Fisher's Least significant difference; A, Net photosynthetic rate; gs, Stomatal conductance; Tr, Transpiration rate; DEGs, Differentially Expressed Genes; FPKM, Fragments per Kilobase of exon per Million fragments mappedapped; EBG, Early biosynthesis genes; LBG, Later biosynthesis genes; OMT, *O*-methyltransferase family protein; GST, Glutathione S-Transferase; PCR, Polymerase Chain Reaction.

MATERIALS AND METHODS

Plant Materials, Phenotypic Characterization, and Leaf Anatomy Observation

In this study, eleven varieties from five different *Brassica* species (*B. rapa*, *B. napus, B. juncea*, *B. oleracea,* and *B. carinata*) were

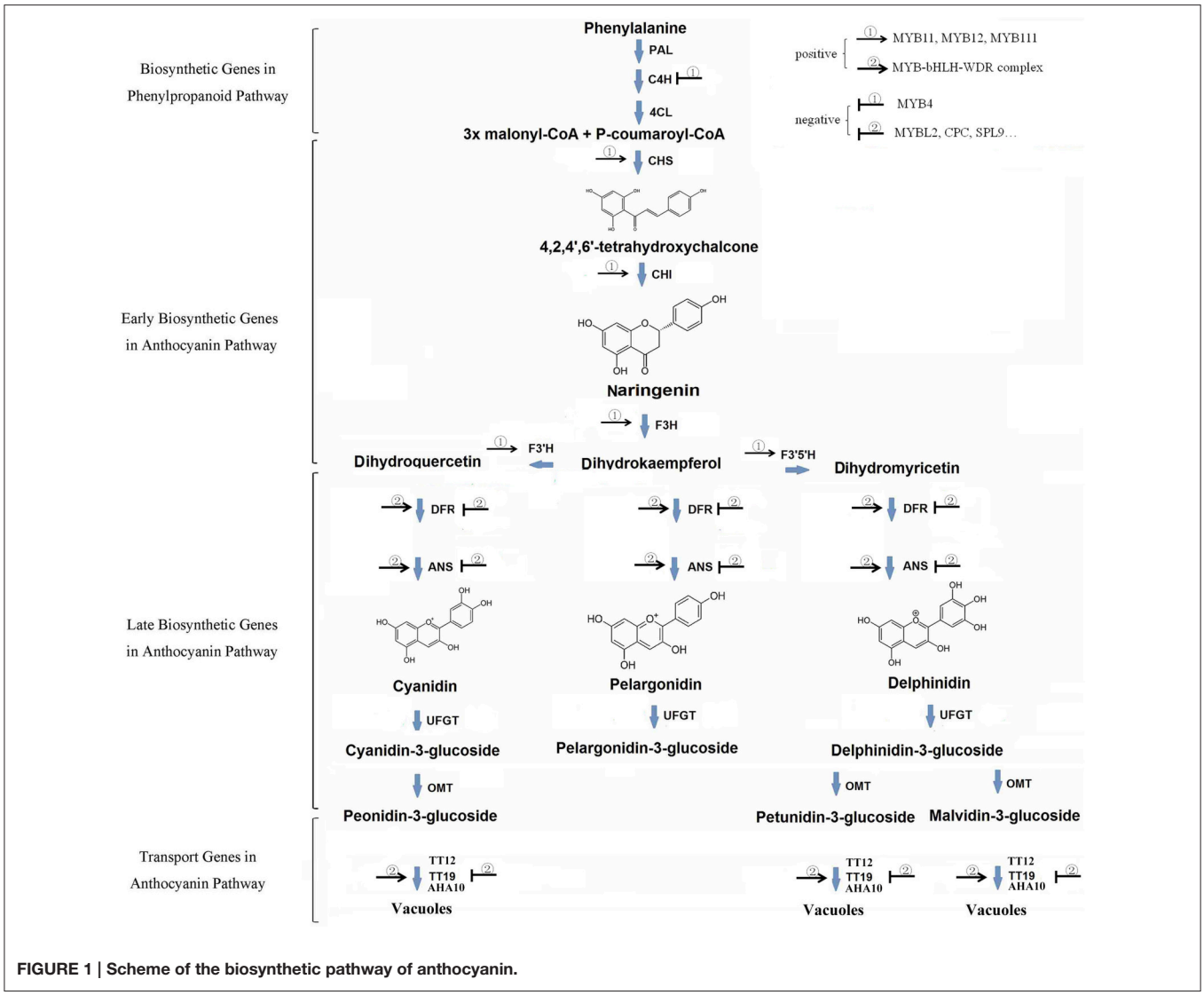


FIGURE 1 | Scheme of the biosynthetic pathway of anthocyanin.

used for phenotype observation, photosynthesis measurement, and RNA-seq analysis. For each species, except for *B. rapa*, one variety with purple and one with green leaves were investigated. For *B. rapa*, two type of varieties with purple and one with green leaves were used (**Table S1**). All varieties were planted in research field of Huazhong Agricultual University, Wuhan, China, during 2013–2014 cropping season. For each species, the purple and green varieties were cultivated side by side. Young leaves at seedling stage were collected at the same day for RNA extraction. Phenotypic comparison was carried out between purple and green for seed endosperms, young leaves, young seedlings of 6-days-old plants, and young plants of 6-weeks-old. To examine different types of cells showing purple pigmentation, Leaves were transversely sectioned by free hand and examined with a Zeiss Axioscope photomicroscope equipped with an MRC digital camera.

Photosynthetic Rate Measurement

Photosynthesis can be measured by photosynthesis measurement systems. These systems measure the rate using an infrared gas analyzer to accesses the input of CO_2 and output of H_2O. In present study, portable photosynthesis machine LI-6400XT (LI-COR Inc., Lincoln, NE, USA) was used for recording photosynthetic rate in all five *Brassica* species. The gas exchange parameters were determined on sunny, windless days from 9:30 to 11:30 a.m. Leaf temperature was controlled at 12°C and photon flux density was maintained at 500 μmol m^{-2} s^{-1}. Net photosynthetic rate (*A*), stomatal conductance (*gs*), and transpiration rate (*Tr*) were recorded on fully expanded leaves of second youngest nodes. The total three readings per treatment were taken from randomly selected plants.

RNA Extraction and mRNA-Seq Libraries Preparation

Young leaves were collected with two biological replicates at seedling stage from the field and immediately frozen in liquid nitrogen and stored at −80°C until use. Total RNA was extracted using TRIzol (Invitrogen) according to the manufacturer's instructions. The quality of purified RNA was initially evaluated on agarose gel and then quantified using

NanoDrop spectrophotometer (Thermo Fisher Scientific, Inc.). The integrity of RNA samples were further evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). The TruSeq TM RNA Sample Preparation Kit (Illumina, Inc.) was then used according to the manufacturer's instructions, to construct cDNA libraries. Concisely, poly-A mRNA was purified and fragmented into short fragments and used as templates for first strand cDNA synthesis. Then DNA polymerase I and RNase H were used to synthesize the second-strand cDNA. Purified short double strand cDNA fragments were connected with adapters (Illumina). Suitable ligated cDNA fragments were selected as templates for the PCR amplification for the finally library construction. Finally, the cDNA libraries were sequenced using Illumina HiSeq 2000 sequencing platform at National Key Laboratory for Crop Genetic Improvement in Huazhong Agricultural University, Wuhan, China.

Processing of Raw Reads and Mapping

Adaptors were removed from the reads firstly. The reads in which unknown bases comprised more than 5% of the total and low quality reads (the percentage of the low quality bases of quality value $\leq$5 is more than 50% in a read) were also removed. The clean reads were aligned to *B. napus* var. Darmor-*bzh* genome accessed from http://www.genoscope.cns.fr/brassicanapus/ allowing up to two mismatches in each segment alignment by Tophat (Trapnell et al., 2010) and Bowtier software. Only those unique mapped reads were used for further analysis.

Assessment of Differentially Expressed Genes (DEGs)

Cufflinks program was used to assemble aligned RNA-Seq reads into transcripts, to estimate their abundances, and to test for differential expression and regulation transcriptome-wide. The gene expression level and the transcripts abundances were calculated using FPKM method. If there were more than one transcript for a gene, the longest one was used to calculate its expression level and coverage. The significance of differential gene expression between the purple and green leaf *Brassica* species was determined using Cuffdiff (adjusted $p \leq 0.001$ and Fold change ≥ 2 as criteria). Heat maps were prepared by HeMI software from the website: http://hemi.biocuckoo.org/faq.php.

Gene Expression Analysis Using Semi qRT-PCR

To evaluate the validity of Illumina analysis and assess the expression profiles in terms of specific mRNA abundance, several genes were selected and detected by Semi qRT-PCR. Reverse transcription was performed by Super Script III Reverse Transcriptase (Invitrogen) and oligo (dT) according to the manufacturer's instructions. Forward and reverse primers were designed by using the Primer 3 software based on conserved sequence of the genes from different species or different copies within the same genome. Sequences of selected genes were obtained from the *B. oleracea* genome database: http://www.ocri-genomics.org/bolbase/index.html, *B. rapa* genome database: http://brassicadb.org/brad/ as well as *B. napus*: http://www.genoscope.cns.fr/brassicanapus/. Beta Actin, as the internal housekeeping gene control, was used to get the bands (25 cycle) using original cDNA. All the cDNA samples were diluted to a concentration which gives same bright bands using the actin primers. Then gene specific primers were used to get different bright bands from different materials (32 cycles). Amplification reactions were performed as the following: an initial denaturation step at 94°C for 5 min, 32 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s, a final extension at 72°C for 10 min and hold at 25°C. The electrophoresis gel run bands were analyzed to verify the specificity of Semi qRT-PCR. All primers used were list in **Table S2**.

RESULTS

Accumulation of Anthocyanins in Different *Brasica* Species

All cultivars used in the study have yellow endosperms and dark seed coat except for BcaP which has a brown to yellow seed coat. Obviously, pigments do not accumulate in endosperms of these *Brassica* cultivars. The purple phenotypes seem to be related primarily with young seedlings, leaves, and young plants (**Figure 2**). The young seedlings of all green cultivars depicted green color except for BjuG with light red color but fade when true leaf was emerged. The young seedlings of BjuP and BolP showed dark red color, but other purple cultivars showed very light or no red color at this stage. During the first few weeks of growth, all the young leaves of green cultivars turned into solid green color, whereas those of purple cultivars displayed purple color in leaves (**Figure 2**). In later development stages, the anthocyanins were accumulated only on leaves in BjuP and BraP1, but in other species, the anthocyanins were found almost in all vegetative tissue of the plant, including petiole, stem, and flower stalk. However, only in BcaP, purple sepals and very little pigments on petals were apparent (data not shown). Under the identical growth conditions, the green plants showed no purple pigmentation in these tissues.

Utilizing free hand sectioning, the mounts of *Brassica* leaf transverse sections were observed. The results revealed that in BraP1, the distribution of purple pigmentation was solidly accumulated only in adaxial epidermal cells (**Figure 3A**). Contrary to BraP1, purple pigmentation was found only in mesophyll cells but not epidermal cells in leaves of BraP2 (**Figure 3B**). Interestingly, almost all palisade and spongy mesophyll cells have pigmentation with decreasing content from edge to center. In BolP, pigmentation was also accumulated in both palisade and spongy mesophyll cells but only in the first layer of the mesophyll cells (**Figure 3C**). In BnaP, purple pigmentation was not found in epidermal and spongy mesophyll cells but very light pigmentation was observed in palisade mesophyll cells (**Figure 3D**). In BjuP, pigmentation was accumulated in both adxial and abxial epidemal cells but almost noting in mesophyll cells (**Figure 3E**). In BcaP, dark purple pigmentation was observed in both adaxial and abaxial epidermal cells as well as in their adjacent one to several layers of palisade or spongy mesophyll cells (**Figure 3F**). In all five

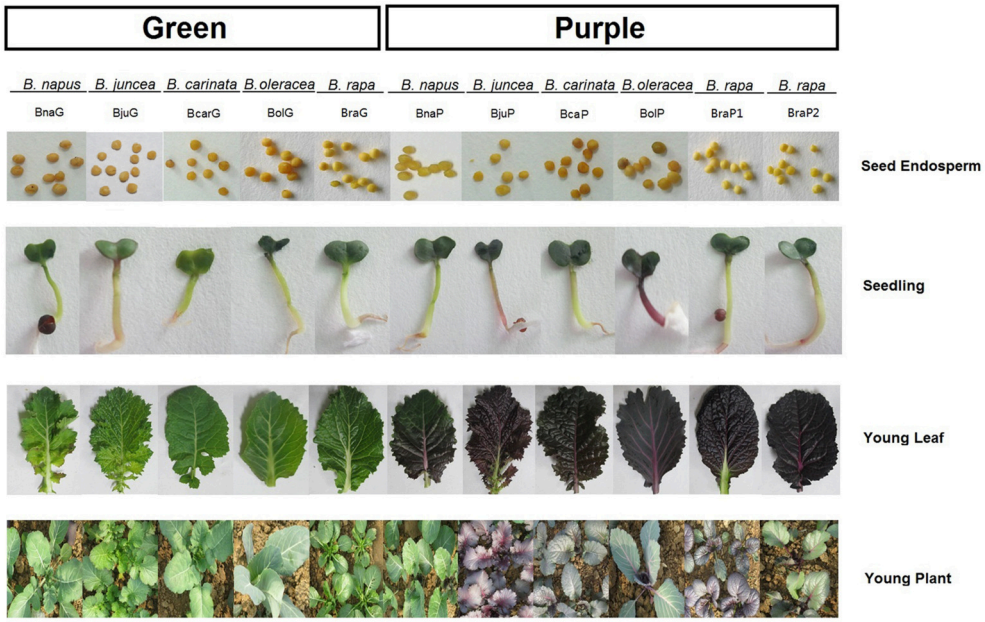


FIGURE 2 | Accumulation of anthocyanin at various tissues/organs of different *Brassica* species.

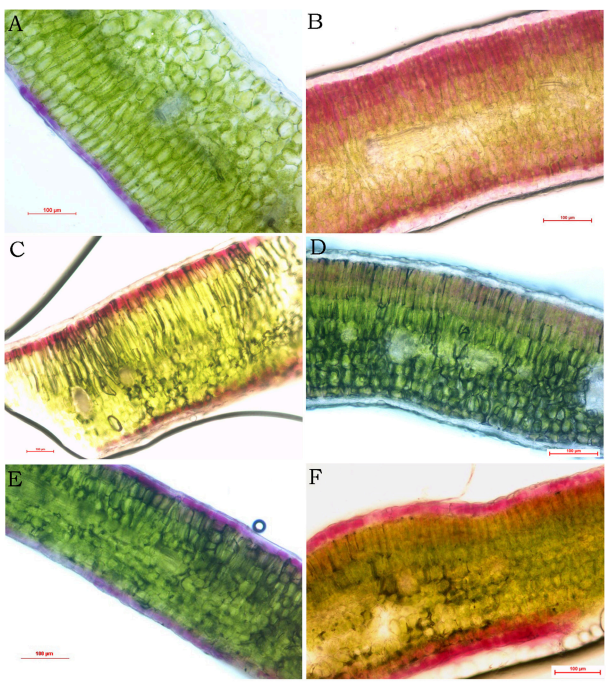


FIGURE 3 | Accumulation of anthocyanin in different part of purple leaves in *Brassica* revealed by hand section. (A–F) purple leaves from *B. rapa* (BraP1), *B. rapa* (BraP2), *B. oleracea* (BolP), *B. napus* (BnaP), *B. juncea* (BjuP), and *B. carinata* (BcaP), respectively.

species, the leaves of all green cultivars did not show purple or red pigmentation in any cell type. In summary, the pattern of anthocyanin accumulation in *Brassica* leaves varies with species. BraP2 and BcaP showed more dark and wide distribution of pigments, while BnaP depicted the lightest purple pigmentation.

Photosynthetic Activity of Leaves with Different Color of Different Species

In order to investigate the potential effect of the anthocyanin accumulation on photosynthesis, we analyzed the photosynthesis-related attributes in all five *Brassica* species. Intraspecific studies revealed that stomatal conductance (*gs*), transpiration rates (*Tr*) as well as rates of photosynthesis were higher in green leaves of *B. juncea*, *B. oleracea,* and *B. rapa* compared with their respective purple leaves (for *B. rapa* only BraP1; **Figure 4**; **Table S3**). Purple leaves of *B. carinara*, *B. rapa*P2 and *B. napus* recorded greater *gs* and *Tr* than their green leaves, however, photosynthetic rate was higher only in BcaP and BraP2 (**Figure 4**; **Table S3**).

Overall Expression of Genes in Anthocyanin Biosynthesis Pathway of *Brassica*

From mRNA sequencing, raw data were obtained from two biological repeats of each purple and green leaves. In total, 498,148,526 raw reads were produced. After removing low quality reads, 455,173,884 clean reads were found (**Table S4**). More than 82.5% reads were mapped in *B. napus* genome from the *B. napus*, but only 37.9% in *B. juncea* and 40.6% in *B.carinata*. The reason might be that *B. carinata* and *B. juncea* have B genome, which is poorly associated with reference genome (A and C). On average, 85.6% mapped reads were unique mapped to the genome. According to each gene expression value (**Supplementary Datasheet S1**), the correlations of two

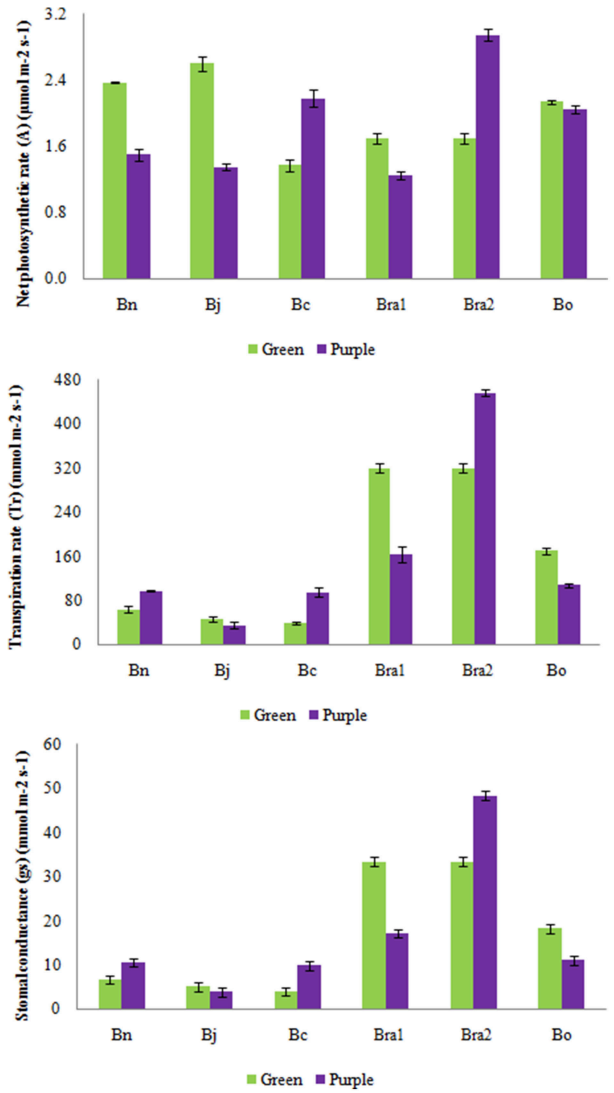


FIGURE 4 | Analysis of leaf stomatal conductance (gs), transpiration rates (Tr) as well as rates of photosynthesis (A) in green and purple leaves of different *Brassica* species. Bn, *B. napus*; Bj, *B. juncea*; Bc, *B. carinata*; Bra1, BraP1 and BraG; Bra2, BraP2 and BraG; Bo, *B. oleracea*.

duplicates were calculated and showed high repeatability in all species (**Table S4**). These results justified the high quality of sequencing data for further analyses.

Previously, 73 genes in *B. rapa* as orthologs of 41 anthocyanin biosynthetic genes in *A. thaliana* have been identified (Guo et al., 2014). Corresponding genes were identified in An and Cn sub-genome of *B. napus* according to http://brassicadb.org/brad/ respectively and their expression value were calculated (**Supplementary Datasheet S2**). It was found that *BrPAL3.1*, *BrPAL3.2*, *BrF3H2*, *BrTTG1.2*, *BrCHS5,* and *BrCHI3* as well as their syntenic genes in C genome are silenced in all materials (**Supplementary Datasheet S2**). To simplify analysis, for each gene in *Arabidopsis*, we added all its pralogous and orthologous genes in each *Brassica* species together for further analysis. Because it was difficult to discriminate *PAP1*, *PAP2*, *MYB113,* and *MYB114* at sequence level in *B. rapa* genome (Guo et al., 2014), *PAP1** was used to represent any of them in followed analysis. On average, 13 genes were expressed at high level (FPKM > 100) and most of which were LBGs, 14 genes expressed at middle level (FPKM = 10–100) and eight genes were expressed at low level (FPKM < 10) in purple leaves. In green leaves, only six genes were expressed at relative high level, including *CHS*, *CHI*, *F3H*, *FLS1*, *PAL1,* and *C4H*. Generally, most of the biosynthesis genes of anthocyanin in purple leaves were expressed at higher level than those in green leaves (**Figure 5**). Based on the expression of each gene, 11 varieties were classified into three main groups. All varieties with purple leaves except for *B. napus* were clustered into same group, while BnaP was grouped with BnaG, BjuG, and BraG (**Figure S1**). BolG and BcaG were divided into independent group. These results clearly showed that *B. carinata* and *B. oleracea* might have similar gene expression pattern while *B. napus* depicted a unique expression pattern of genes in anthocyanin synthesis.

Differential Expression of Structural Genes of Anthocyanin Biosynthesis Pathway Between Purple and Green Leaves

Among eight genes in phenylpropopanoid metabolism, *4CL5* was expressed at very low level in all varieties and almost no difference was found between green and purple leaves, except for *B. oleracea*, where it was expressed at a little higher level in purple leaves (**Figure 6**). Other genes were highly expressed in purple leaves of *B. carinata* and *B. oleracea* and have higher expression level than those in respectively green leaves. In *B. rapa* and *B. juncea*, all other genes but *4CL3* were expressed at higher level in purple leaves than those in green leaves. However, only *PAL4* and *4CL2* were expressed at higher level in purple leaves than those in green leaves of *B. napus* (**Figure 6**). For EBGs, *CHS*, *CHI,* and *F3H* were expressed at much higher level in purple leaves of *B. carinata* and *B. oleracea* as well as a little higher level in BraP1, but showed similar level in green and purple leaves or lower level in purple leaves of other species (**Figure 7**). F3′H showed lowest expression in *B. napus* but much higher expression in *B. carinata* and *B. oleracea* in purple leaves. *FLS1* was expressed at lower level in purple leaves of all species except for *B. carinata*, where it showed a little higher expression. *FLS2*, *FL3*, and *FL4* were silenced or expressed at very low level (FPKM < 10) in all materials except for in *B. oleracea*. *FLS2* in *B. napus* and *B. oleracea*, *FL3* in *B. juncea*, and *FLS4* in *B. oleracea* showed higher expression in purple leaves. *FLS5* was down regulated in purples of *B. napus*, BraP1, and BraP2, but a little up-regulated in purple leaves of *B. oleracea*, *B. juncea,* and *B. oleracea*. *DFR*, *ANS* as well as *TT19* were greatly up-regulated in purple leaves of all varieties. Meanwhile, *UGT75C1*, *UGT79D2,* and *UGT78B1* depicted much higher expression in all purple leaves except for *UGT79D2* in *B. napus* and *UGT78B1* in *B. oleracea* where it was little down-regulated. In summary, EBGs of anthocyanin showed no prominent differences between green and purple leaves of *B. napus*, *B. juncea,* and *B. rapa*, but were highly up-regulated in purple leaves of *B. carinata* and *B. oleracea*. However,

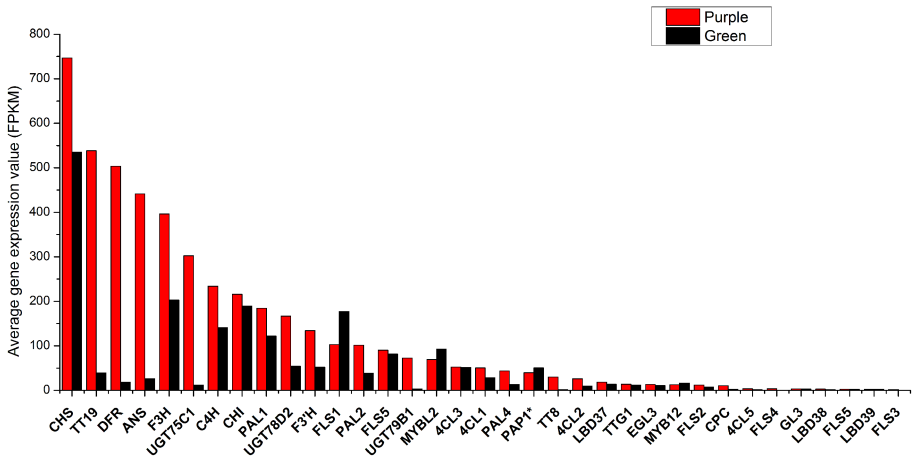


FIGURE 5 | Average total expression value (FPKM) of each gene in anthocyanin biosynthesis pathway in purple and green leaves of different *Brassica* species.

almost all LBGs, especially for *DFR*, *ANS,* and *TT19*, were highly up-regulated in all purple leaves.

Differential Expression of Regulating Genes of Anthocyanin Biosynthesis Pathway Between Purple and Green Leaves

It is well known that a series of transcript factor (TFs) play key roles in regulation of structural genes expression, and resulting in the accumulation of anthocyanin. In this study, all the TFs were expressed at relative low level in comparison with most of the structural genes. *MYB12* and MYB111 were only up-regulated in purple leaves of *B. carinata*, while down-regulated or showed no prominent differences between green and purple leaves of other varieties. *PAP1** (*PAP2*, *MYB113,* and *MYB114*), *TT8*, *TTG1,* and *EGL3* and *GL3* are the core proteins for MBW complex. It was found that *PAP1* was expressed at a little higher level in all purple leaves except for *B. juncea* and *B. carinata* (**Figure 8**). *GL3* and *EGL3* were expressed at very low level in all species except for *B. carinata* and *B. oleracea*, in which *EGL3* depicted relative higher expression in purple leaves. The *TT8* was the only positive TF which up-regulated in all purple leaves, especially in BraP2, *B. juncea*, *B. carinata,* and *B. oleracea* (**Figure 8**). For negative regulators, CPC was up-regulated at different degree in all purple leaves expect for *B. napus,* where it was down-regulated. MYBL2 was expressed at higher level in all leaves and was down regulated in purple leaves of *B. napus*, *B.carinata*, *B. oleracea,* and *B. juncea* but up-regulated in two *B. rapa*. *LBD 37*, *LBD38,* and *LBD39* were expressed at very low level. *LBD37* were little higher in purple leaves of all species except for *B. oleracea*. Although the expression pattern of *LBD* 38 and *LBD 39* was also similar in other purple and green leaves, *LBD38* in BcaG, *LBD39* in BcaG and BraG was silenced (**Figure 8**; **Supplementary Datasheet S3**).

In order to identify potential key regulators tightly related to anthocyanin accumulation in *Brassica*, we performed clustering analysis of all genes and varieties based on gene expression changes (fold change) between purple and green leaves (**Figure 9**). While two *B. rapa*, *B. napus* and *B. juncea* were clustered separately into two groups, *B. oleracea* and *B. carinata* occupied relative independent branches. This result indicated that two types of *B. rapa*, *B. juncea* and *B.napus* might have similar mechanism respectively for anthocyanin accumulation in leaves. Most of structural genes were clearly clusters into two major subgroups, one included genes involved in phenylpropopanoid pathway and EBGs and the other included all of LBGs (**Figures 1**, **9**). Meanwhile, transcription factors involved in MBW complex were grouped with those LBGs except for GL3 and EGL3. Interestingly, some of the transcripts were grouped into independent subgroups, for example, EGL3, MYB11 and MYB111, CPC and LBD38, *PAP1*, *LBD38, TTG1,* and *MYBL2*. These results are well consist with that there are complex mutual regulating relationships between different transcript factors in anthocyanin synthesis (Petroni and Tonelli, 2011). Prominently, *DFR*, *ANS*, *TT19*, *UGT75C1*, *UGT79B1,* and only transcript factor, *TT8*, are cluster together prominently, indicating that expression changes of *TT8* has major effect on the expression of LBGs for anthocyanin accumulation in *Brassica* leaves.

RT-PCR Validation of Gene Expression in Anthocyanin Biosynthesis Pathway

To verify the relationships of gene expression revealed by RNA-seq analysis between green and purple leaves, a total of nine genes from anthocyanin biosynthetic pathway were chosen to perform semi-quantitative RT-PCR. Because these primers were designed based on the conserved sequence of each gene, the results will reflect the total expression values other than from specific copies. The relative transcript levels in all purple and green *Brassica* species were compared with those in RNA-seq data. In *B. oleracea*, qRT-PCR analysis showed that all genes but *PAL1* has the same expression trends as those in RNA-seq data (**Figure S2A**; **Figures 6**–**8**). In other four species,

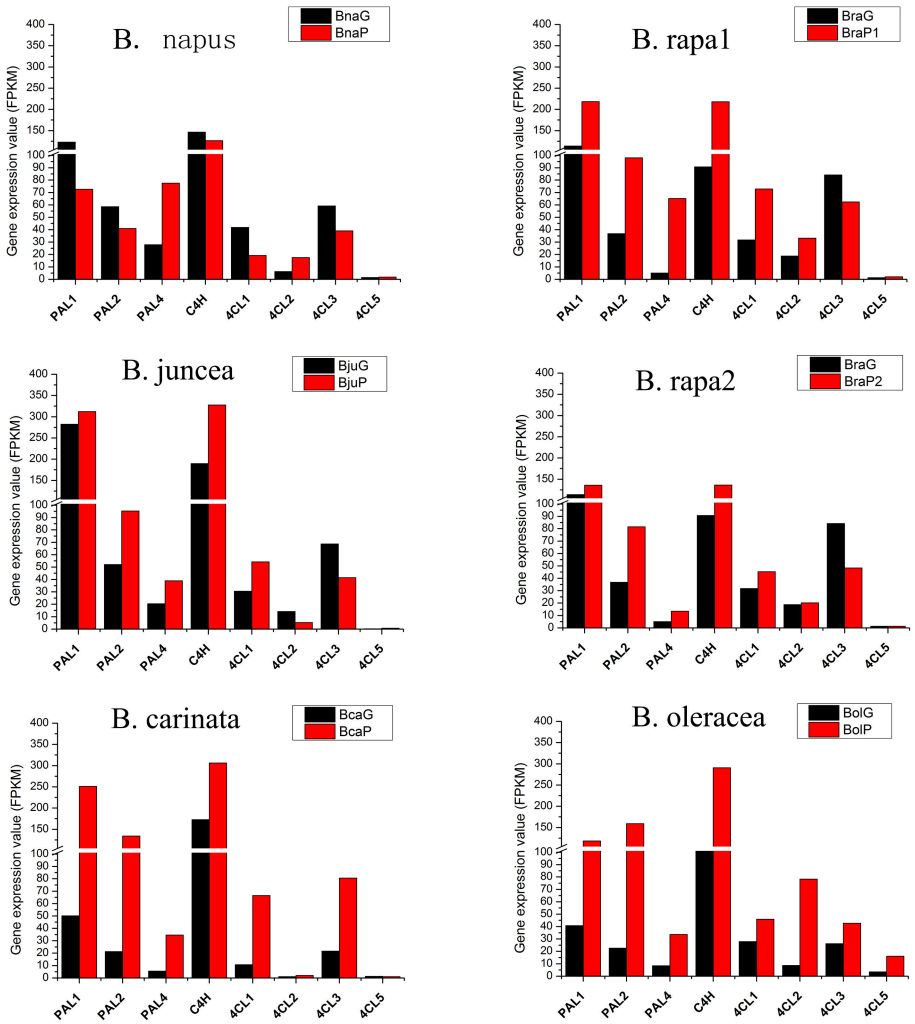


FIGURE 6 | Total expression value of structural genes in phenylpropanoid biosynthesis pathway in green and purple leaves of different *Brassica* species.

seven genes were performed RT-PCR analysis while *PAP1* and *CHS* cannot work well. Most of these genes also have the same expression trends as those in RNA-seq data, especially for *TT8* and *MYBL2* (**Figure S2B**; **Figures 6–8**). These results further confirmed the reliability of RNA-seq data in present study.

Co-DEGs Analysis Between Purple and Green Leaves of Different *Brassica* Species

In order to analyze the potential effect of anthocyanin accumulation on other gene expression, we firstly analyzed the differentially expressed genes between purple and green varieties and identified same DEGs by comparing the varieties with purple and green leaves. It was found that *B. oleracea* had the largest number of DEGs (6582) while *B. rapa 2* has the smallest number of DEGs (1057) (**Supplementary Datasheet S4**; **Table S5**). In order to analysis common DEGs in all species, we converted gene ID of *B. napus* into gene ID of *Arabidopsis*, because 11 varieties are belonging to five species with different genome. Finally, 21 common DEGs were found in all comparisons. Most of these genes have roles in photosynthesis, anthocyanin synthesis, and ribosome components (**Table S6**). We then calculated total expression value of these genes in each variety (**Supplementary Datasheet S5**). It was noteworthy, that *PLG5* which encodes a cytosolic 6-phosphogluconolactonase and thought to be involved in the oxidative pentose-phosphate pathway (OPPP), was highly expressed in all purple leaves along with *TT19*, *ANS*, and *DFR*. Meanwhile, four genes are co-down regulated in varieties with purple leaves, including *FTS8*, *GOX1*, *GLN1*, and *COR27* (**Table S6**). *FTS8* encode FtsH protease that is localized to the chloroplast. GOX1 encodes a glycolate oxidase that is the key genes in light respiration and modulates reactive oxygen species-mediated signal transduction during non-host resistance. *GLN;1.4* encodes a cytosolic glutamine synthetase, and takes part in assimilation process of the ammonia produced by the light respiration and the reduction of nitric

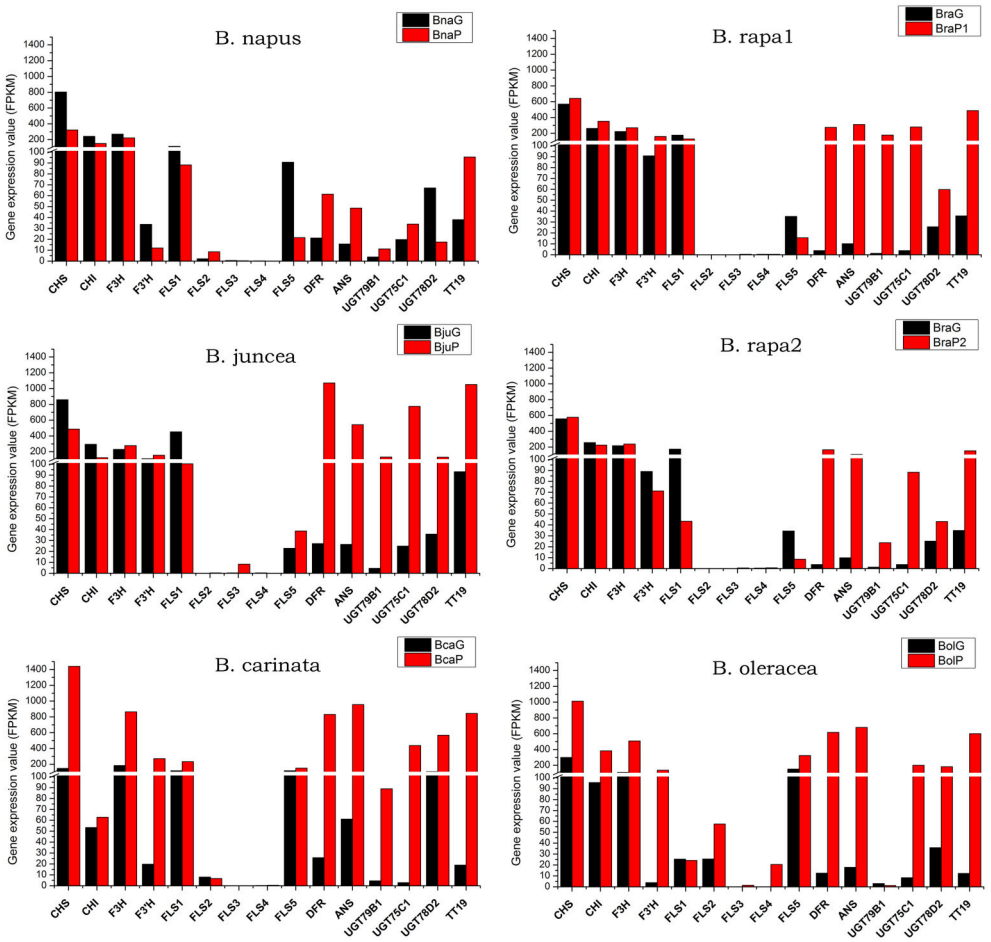


FIGURE 7 | Total expression value of structural genes in anthocyanin biosynthesis pathway in green and purple leaves of different *Brassica* species.

acid. However, there is no function information of *COR27* till date.

DISCUSSION

With the rapid decrease in cost, trascriptome analysis using RNA-seq technology have become one of the most frequent and reliable methods for gene identification, genome evolution, developmental regulation, and genetic mapping studies. The available genome for *B. rapa*, *B. oleracea* and *B. napus* greatly facilitated the transcriptome analysis in *Brassica* species, especially for homeologous genes identification (Wang et al., 2011; Chalhoub et al., 2014; Liu et al., 2014; Parkin et al., 2014). Here, we performed comparative transcriptome analysis between different varieties belonging to five species with purple and green leaves. Using the *B. napus* as reference genome, we emphasized on the expression of the genes related with anthocyanin synthesis in *B. rapa* (Guo et al., 2014). Although each variety with purple and green leaves has different genetic background, we are focusing only on those similar results in different comparisons which will provide us credible conclusion.

Leaf Pigmentation and Genome Relationships in *Brassica*

Variations in leaf pigmentation are common in *Brassica* especially those used for vegetables, for example, red cabbage and purple cauliflower in *B. oleracea* and "Hongshancaitai" in *B. rapa* in China, but limited information is available in *B. nigra*. Because *Brassica* tetraploids are evolved from the hybrids between pair of diploids, it is reasonable to suspect that the variation of leaf pigmentation in tetraploids might be due to variations in corresponding parental diploids. Clustering analysis based on the expression of genes for anthocyanin synthesis in different species also reflected their genome relationships. The varieties with purple leaves were in same group except for *B. napus*. Meanwhile, BraP1, BraP2, and BjuP were occurred in the same group and BolP and BcaP were classified into other group (**Figure S2**). However, free hand section analysis showed that the distribution of pigmentation in leaves varies with species (**Figure 3**). The pattern of pigmentation in leaves of tetraploids was different from those in diploids. It looks that the variation in tetradploids was occurred independently from those diploids. An excellent example is that the *B. napus* with purple leaves used

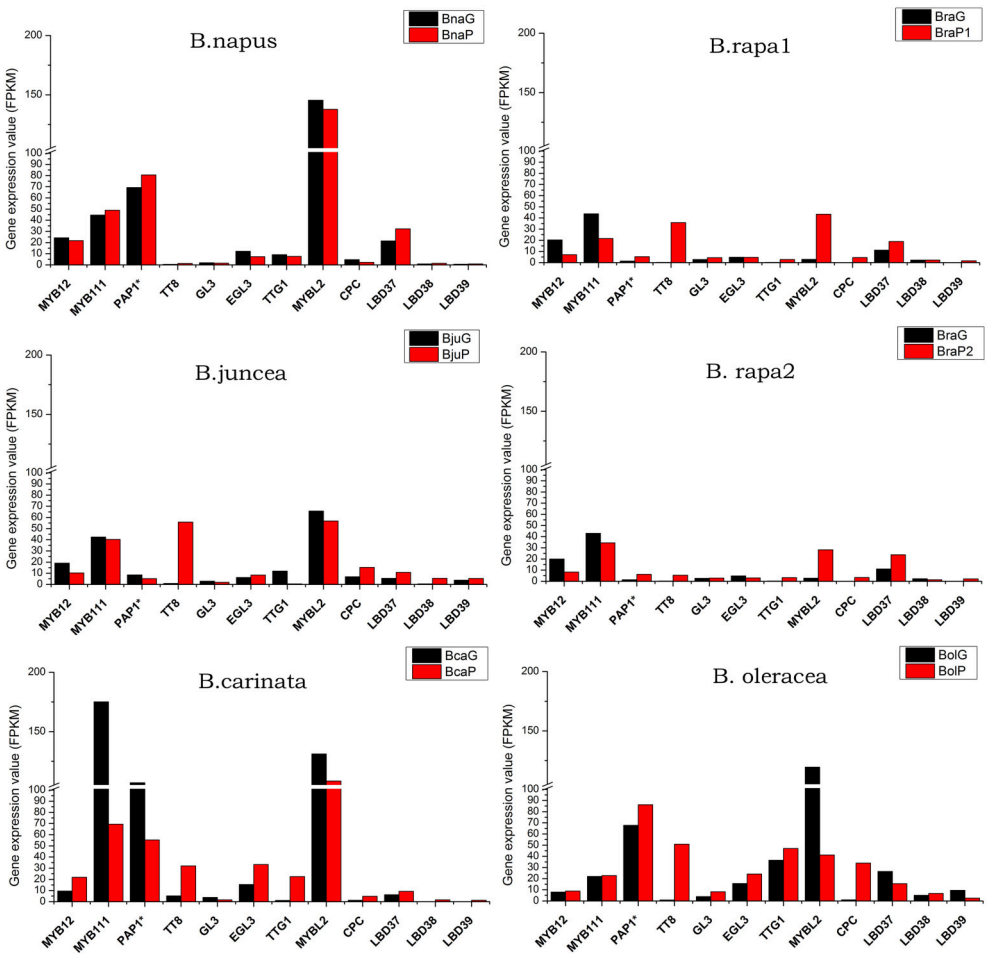


FIGURE 8 | Total expression value of regulator genes in anthocyanin biosynthesis pathway in green and purple leaves of different *Brassica* species. PAP1* represent any of PAP1, PAP2, MYB113, and MYB114.

here is one of the progenies from wide hybridization between *B. napus* and *O. violaceus* (Ge et al., unpublished data). This *B. napus* have a different distribution of pigmentation in leaves and a very special expression pattern of those genes response for anthocyanin biosynthesis (**Figures 2**,**3**,**9**; **Figure S1**).

TT8 Plays Key Roles in Regulating Leaf Pigmentation in *Brassica*

Many studies had been done to find key genes response for pigmentation in *Brassica* species at gene expression level. *BoMYB2* and *BrTT8* were two regulators identified by mapping based cloning for pigmentation in curds of *cauliflower* (*Brassica oleracea* var *botrytis*) and seeds of *Brassica rapa,* respectively till now (Chiu et al., 2010; Li et al., 2012). In this study, comparative transcriptome analysis between paired green and purple leaves of 11 varieties belonging to five different species clearly showed that the most of anthocyanin biosynthesis genes were greatly up-regulated, especially for those LBGs (**Figure 7**). These results indicated that anthocyanin accumulation in different part of leaves of different species might results from the variation of regulators involved in MBW complex which regulating LBGs. While other components of MBW are up or down regulated in purple leaves, *TT8* was up-regulated along with LBGs in all purple leaves. Particularly, the expression values of *TT8* generally appear positive relationships with the anthyocyanin content (**Figures 2**, **3**, **8**). Our results were in accordance with previous studies investigated the gene expression by RT-PCR and suggest that TT8 was highly up-regulated in all *Brassica* with purple tissues and appeared to be a key candidate gene (Yuan et al., 2009; Xie et al., 2014; Zhang et al., 2014, 2015c; Ahmed et al., 2015).

In *Arabidopsis*, the main components of the MBW complexes are WD40, bHLH and MYB transcription factors. WD40 transcription factor is encoded by *TTG1*, bHLH are encoded by *TT8, GL3* and *EGL3* and MYB are encoded by *PAP1*, *PAP2*, *MYB113,* and *MYB114* (Borevitz et al., 2000; Ramsay and Glover, 2005; Gonzalez et al., 2008). It was found that in *Arabidopsis* TT8 was required for the full transcriptional activation of late biosynthesis genes (Nesi et al., 2000) although it exhibits partially functional redundancy with *GL3* and *EGL3* (Zhang et al., 2003).

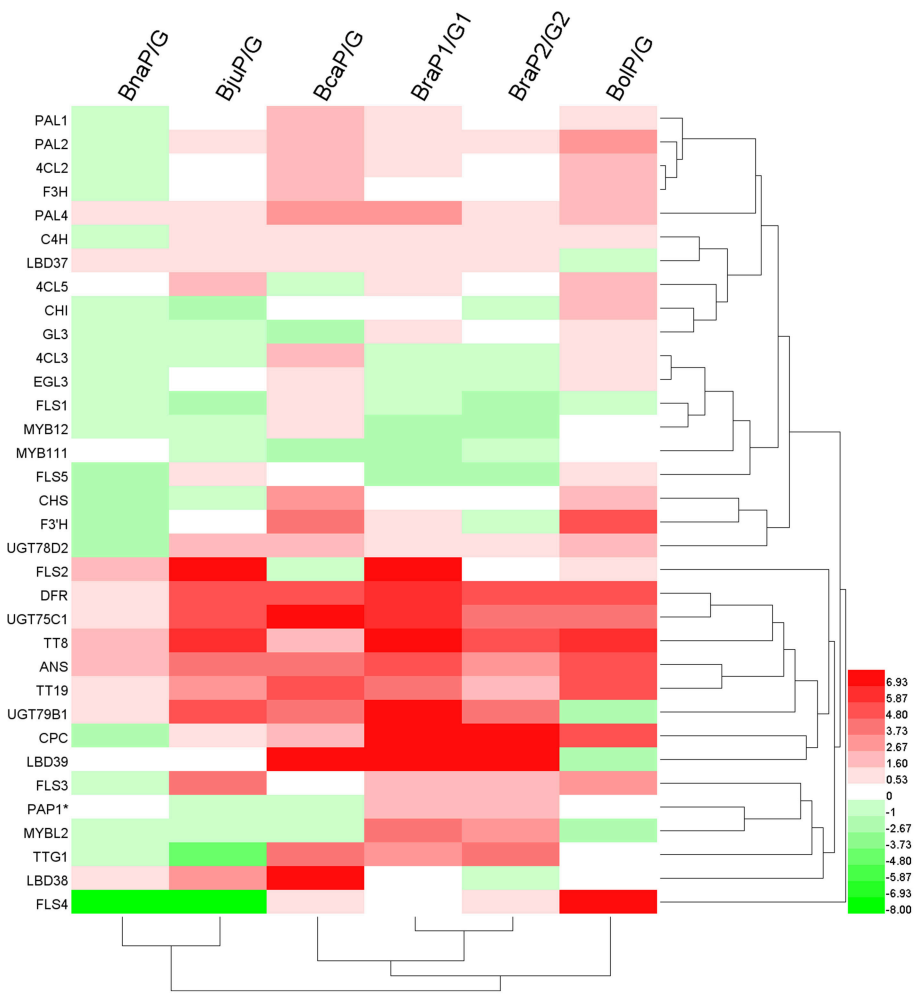


FIGURE 9 | Clustering analysis of all genes in anthocyanin biosynthesis pathway based on the fold changes between green and purple leaves. Color bar: Log_2 (fold changes). PAP1* represent any of PAP1, PAP2, MYB113, and MYB114. Note: because some of genes were silenced (expression value = 0; **Datasheet S3**) in one of green or purple leaves, fold changes were set by hand according to the difference of each comparison.

GL3 and *EGL3* contribute equally to the activation of *F3′H*, but *EGL3* appears more predominant in activation of *DFR* and *ANS* in Arabidopsis seedling pigmentation (Gonzalez et al., 2008). Here, *GL3* and *EGL3* expressed at very low level and have far clustering relationships with late biosynthesis genes (**Figure 9**). *TT8* thus might the key bHLH factor in MBW complex for leaves pigmentation in *Brassica* species. A complex regulation network among these regulators had been described (Zhang et al., 2003). *TT8* appears to be positively regulated by an MBW complex including the WD40 (TTG1), the MYB (TT2, PAP1/PAP2/MYB113/MYB114), and the bHLH itself or GL3/EGL3 and negatively regulated by MYBL2. Meanwhile, it also positively regulated TT2, PAP1/PAP2/MYB113/MYB114, GL3/EGL3, and MYBL2 (Baudry et al., 2004; Gonzalez et al., 2008; Petroni and Tonelli, 2011). Two modules were found to specifically drive *TT8* promoter activity by differentially integrating the signals issued from different regulators in a spatiotemporal manner which involves at least six different MBW complexes (Xu et al., 2013). These complex regulation relationships might make it difficult to indentify linear dependence between any two regulators as found here. For example, while *TT8* was upregulated in all purple leaves, *MYBL2* is upregulated in two *B. rapa* but down regulated in other purple leaves (**Figure 8**).

Potential Function of Anthocyanin Accumulation in Leaves in *Brassica*

Although anthocyanins exist commonly in different species, their function in plant environment interactions remain highly contested. Till date, four putative functions of anthocyanins were proposed in plant development including (1) sunscreens and antioxidants, (2) mediators of reactive oxygen species (ROS)-induced signaling cascades; (3) anthocyanins may serve as metal-chelating agents under conditions of excess edaphic

metal ions; (4) delayers of leaf senescence, especially in plants growing under nutrient deficiency (Landi et al., 2015). The first two types of functions were thought to be tightly related to photosynthesis. Because anthocyanin absorb strongly in the blue-green waveband, thus effectively reduce the wavelengths and intensity of light available to be used for photosynthesis (Nicole, 2009). Meanwhile, under high-light condition the down-regulation of internal light by anthocyanin plays key roles in photoprotection (Nicole, 2009). Additionally, as strong antioxidant, anthocyanin could protect tissues against radical oxygen species (ROS) generated in the chloroplast (Manetas, 2006).

The accumulation of anthocyanins within leaf tissues varied significantly among species or among different varieties of same species. Usually, anthocyanins are accumulated and stored in cell vacuoles, in or just below the adaxial epidermis, but sometimes these pigments entangled in photo-protection, and also accumulate in cell vacuoles of the abaxial epidermis, palisade and spongy mesophyll cells of leaves (Hatier and Gould, 2009). Our results showed that photosynthetic rate was higher in purple leaves of *B. carinata* and B. raP2 but lower in purple leaves of other species (**Figure 4**). Meanwhile, hand section analysis showed that high content of anthocyanin was absolutely in mesophyll of purple leaves in *B. rapa* P2 and both in epidemical and mesophyll cells in *B. carianta* (**Figures 2**, **3**). Although anthocyanin was also distributed in mesophyll cells in purple leaves of *B. napus* and *B. oleracea,* but either the content is very lower in *B. napus* or only in limited layer of mesophyll cells in *B. oleracea*. It looks that high content of anthocyanin located in mesophyll might promote photosynthesis. However, this needs more evidence to make confirmation because the materials used here have very different genetic background. And many studies in other species had been shown that leaves with accumulation of anthocyanin have less photosynthesise than green leaves (Gould et al., 2002). So, comparison analysis between purple and green leaves with same genetic background would help to explore this question in *Brassica* in future (Tohge et al., 2005).

Interestingly, we found that 21 same DEGs between green and purple leave of all species, which potentially results from the accumulation of anthocyanin in purple leaves (**Figure 9**). These included *PLG5*, *FTS8, GOX1*, *GLN1;4,* and *COR27* which is co-upregulated or co-downregulated in all purple leaves. PLG5 is a key enzyme in oxidative pentose-phosphate pathway (OPPP) which produces many intermediate products including phenylalanine, the precursor for anthocyanin synthesis. The upregulation of *PLG5* in purple leaves are well with that anthocyanin accumulation there. FtsH belongs to subfamily ATPases Associated with diverse cellular Activities (AAA+). In plants, FtsH exists as a heterocomplex comprising isomers of two types: FtsH5/FtsH1 (Type A) and FtsH2/FtsH8 (Type B) (Yu et al., 2004, 2005; Zaltsman et al., 2005a). A series of studies indicated that thylakoid-embedded FtsH degrades photo-damaged D1 reaction center proteins in the photosystem II (PSII) repair cycle (Kato and Sakamoto, 2010; Nixon et al., 2010) and to control chloroplast development (Lindahl et al., 2000; Bailey et al., 2002; Sakamoto, 2003; Zaltsman et al., 2005a,b). Because studies had showed that foliar anthocyanins can protect chloroplasts from the adverse effects of excess light (Pietrini et al., 2002; Steyn et al., 2002; Close and Beadle, 2003; Hughes and Smith, 2007; Gould et al., 2010), anthocyanins accumulated in *Brassica* purple leaves might absorb a fraction of the yellow/green and ultraviolet wavelengths, and thus reduce the damage to PSII, and in particular that related to D1 repair and the oxygen evolving complex (OEC; Landi et al., 2015). These might lead to the down-regulation of *FstH8* in purple leaves. GOX1 and GLN1;4 are two important enzyme in plant photorespiration. *GOX1* encodes a glycolate oxidase catalyzes the conversion of glycolate into glyoxylate during photorespiration with concomitant production of H_2O_2 in peroxisomes (Foyer and Noctor, 2009). *GLN1;4* encodes a cytosolic glutamine synthetase, take part in assimilation process of the ammonia produced by the light respiration in chloroplast and the reduction of nitric acid. Simultaneously down-regulated of *GOX1* and *GLN1;4* in all purple leaves indicated that the photorespiration might be repressed in purple leaves in *Brassica.* Similarity, in the *Arabidopsis chi/f3′h* mutant that does not accumulate anthocyanins, the expression levels of *GOX1* were lower than in the wild type (Zhou et al., 2013). In summary, although we did not find direct relationships between the photosynthesis rate and the accumulation of anthocyanin in leaves of *Brassica* species, it looks that purple pigments play roles in reducing the damage to PSII and repressing photorespiration which were reflected at gene expression level.

CONCLUSION

We performed comprehensive gene expression analysis related to anthocyanin biosynthesis along with phenotype and physiological observations in different *Brassica* species with purple and green leaves. Our results depicted that anthocyanin was accumulated in different parts of the leaves in different *Brassica* species, which might be due to the variations in core components of MBW complex. Particularly, TT8 might play pivotal role for the regulation of leaf pigmentation because it was the only up-regulated transcript factor in all purple leaves. Meanwhile, genes involved in OPPP pathway and light respiration were also found co-up or co-down regulated accompanied with the accumulation of anthocyanin in purple leaves which indicated the possible interaction between anthocyanin and photosynthesis.

AUTHOR CONTRIBUTIONS

The study was conceived by XG. MM prepared the plant materials and performed the experiments. QZ performed sequencing. MM, QP, and XG performed the bioinformatics analysis. MM, XG, and ZL prepared the manuscript. All authors contributed to revising the manuscript. All authors had read and approved the final manuscript.

AVAILABILITY OF SUPPORTING DATA

All the sequencing data used in this research has been submitted to public database NCBI under PRJNA298501 (accession number: SRP064721).

ACKNOWLEDGMENTS

We thank Saddam Hussain for revision of the manuscript. This work was supported by Fundamental Research Funds for the Central Universities (2662015PY053).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00311

Figure S1 | Clustering analysis of all genes in anthocyanin biosynthetic pathway based on total expression value between green and purple leaves. PAP1* represent any of PAP1, PAP2, MYB113, and MYB114.

Figure S2 | RT-PCR analysis of some genes in anthocyanin biosynthesis pathway.

Table S1 | All materials used in this study.

Table S2 | Primers used for RT-PCR analysis in this study.

Table S3 | Analysis of leaf stomatal conductance (gs), transpiration rates (Tr) as well as rates of photosynthesis (A) in green and purple leaves of different *Brassica* species.

Table S4 | Number of raw and clean reads, their mapping and repeatability analysis.

Table S5 | The numbers of DEGs between green and purple leaves of different species.

Table S6 | Functional annotation of co-DEGs between purple and green leaves of different species.

Supplementary Data Sheet S1 | Expression value of each gene identified in different samples in different *Brassica* species.

Supplementary Data Sheet S2 | Expression value of each gene in anthocyanin biosynthesis pathway.

Supplementary Data Sheet S3 | Total expression value and fold changes between green and purple leaves of genes in biosynthesis pathway of anthocyanin.

Supplementary Data Sheet S4 | Analysis of differentially expressed genes between green and purple leaves.

Supplementary Data Sheet S5 | Total expression value of 21 co-DEGs between purple and green leaves in different *Brassica* species.

REFERENCES

Ahmed, N. U., Park, J. I., Jung, H. J., Hur, Y., and Nou, I. S. (2015). Anthocyanin biosynthesis for cold and freezing stress tolerance and desirable color in *Brassica rapa*. *Funct. Integr. Genomics* 15, 383–394. doi: 10.1007/s10142-014-0427-7

Bailey, S., Thompson, E., Nixon, P. J., Horton, P., Mullineaux, C. W., Robinson, C., et al. (2002). A critical role for the Var2 FtsH homologue of *Arabidopsis thaliana* in the photosystem II repair cycle *in vivo*. *J. Biol. Chem.* 277, 2006–2011. doi: 10.1074/jbc.M105878200

Bassolino, L., Zhang, Y., Schoonbeek, H. J., Kiferle, C., Perata, P., and Martin, C. (2013). Accumulation of anthocyanins in tomato skin extends shelf life. *New Phytol.* 200, 650–655. doi: 10.1111/nph.12524

Baudry, A., Heim, M. A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J.* 39, 366–380. doi: 10.1111/j.1365-313X.2004.02138.x

Bohm, B. A. (1998). *Introduction to Flavonoids, Vol. 2*. Amsterdam: Hardwood Academic Publishers.

Borevitz, J. O., Xia, Y., Blount, J., Dixon, R. A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12, 2383–2394. doi: 10.1105/tpc.12.12.2383

Butelli, E., Titta, L., Giorgio, M., Mock, H.-P., Matros, A., Peterek, S., et al. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat. Biotechnol.* 26, 1301–1308. doi: 10.1038/nbt.1506

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* 70, 1–9. doi: 10.1111/j.1751-1097.1999.tb01944.x

Chiu, L. W., Zhou, X., Burke, S., Wu, X., Prior, R. L., and Li, L. (2010). The purple cauliflower arises from activation of a MYB transcription factor. *Plant Physiol.* 154, 1470–1480. doi: 10.1104/pp.110.164160

Close, D. C., and Beadle, C. L. (2003). The ecophysiology of foliar anthocyanin. *Bot. Rev.* 69, 149–161. doi: 10.1663/0006-8101(2003)069[0149:TEOFA]2.0.CO;2

Dao, T. T. H., Linthorst, H. J. M., and Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochem. Rev.* 10, 397–412. doi: 10.1007/s11101-011-9211-7

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., et al. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *Plant J.* 55, 940–953. doi: 10.1111/j.1365-313X.2008.03564.x

Foyer, C. H., and Noctor, G. (2009). Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxid. Redox Signal.* 11, 861–905. doi: 10.1089/ars.2008.2177

Glover, B. J., and Martin, C. (2012). Anthocyanins. *Curr. Biol.* 22, R147–R150. doi: 10.1016/j.cub.2012.01.021

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 53, 814–827. doi: 10.1111/j.1365-313X.2007.03373.x

Gou, J. Y., Felippes, F. F., Liu, C. J., Weigel, D., and Wang, J. W. (2011). Negative regulation of anthocyanin biosynthesis in Arabidopsis by a miR156-targeted SPL transcription factor. *Plant Cell* 23, 1512–1522. doi: 10.1105/tpc.111.084525

Gould, K. S., Dudle, D. A., and Neufeld, H. S. (2010). Why some stems are red: cauline anthocyanins shield photosystem II against high light stress. *J. Exp. Bot.* 61, 2707–2717. doi: 10.1093/jxb/erq106

Gould, K. S., Vogelmann, T. C., Han, T., and Clearwater, M. J. (2002). Profiles of photosynthesis within red and green leaves of *Quintinia serrata*. *Physiol. Plantarum* 116, 127–133. doi: 10.1034/j.1399-3054.2002.1160116.x

Guo, N., Cheng, F., Wu, J., Liu, B., Zheng, S., Liang, J., et al. (2014). Anthocyanin biosynthetic genes in *Brassica rapa*. *BMC Genomics* 15:426. doi: 10.1186/1471-2164-15-426

Harborne, J. B., and Grayer, R. J. (1994). "Flavonoids and insects," in *The Flavonoids, Advances in Research Since 1986,* ed J. B. Harborne (London: Chapman & Hall), 589–618.

Hatier, J. H. B., and Gould, K. S. (2009). "Anthocyanin function in vegetative organs," in *Anthocyanins: Biosynthesis, Functions, and Applications*, eds K. S. Gould, K. M. Davies, and C. Winefield (New York, NY: Springer), 1–19.

Hatier, J. H., and Gould, K. S. (2008). Foliar anthocyanins as modulators of stress signals. *J. Theor. Biol.* 253, 625–627. doi: 10.1016/j.jtbi.2008.04.018

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J. Exp. Bot.* 62, 2465–2483. doi: 10.1093/jxb/erq442

Hou, D. X. (2003). Potential mechanisms of cancer chemoprevention by anthocyanins. *Curr. Mol. Med.* 3, 149–159. doi: 10.2174/1566524033361555

Hughes, N. M., Neufeld, H. S., and Burkey, K. O. (2005). Functional role of anthocyanins in high-light winter leaves of the evergreen herb *Galax urceolata*. *New Phytol.* 168, 575–587. doi: 10.1111/j.1469-8137.2005.01546.x

Hughes, N. M., and Smith, W. K. (2007). Attenuation of incident light in *Galax urceolata* (Diapensiaceae): concerted influence of adaxial and abaxial anthocyanic layers on photoprotection. *Am. J. Bot.* 94, 784–790. doi: 10.3732/ajb.94.5.784

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., et al. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J.* 19, 6150–6161. doi: 10.1093/emboj/19.22.6150

Kato, Y., and Sakamoto, W. (2010). New insights into the types and function of proteases in plastids. *Int. Rev. Cell Mol. Biol.* 280, 185–218. doi: 10.1016/S1937-6448(10)80004-8

Kerio, L. C. (2011). The functional significance of anthocyanins in leaves: a review. *Tea* 32, 46–52.

Landi, M., Tattini, M., and Gould, K. S. (2015). Multiple functional roles of anthocyanins in plant-environment interactions. *Environ. Exp. Bot.* 119, 4–17. doi: 10.1016/j.envexpbot.2015.05.012

Lea, U. S., Slimestad, R., Smedvig, P., and Lillo, C. (2007). Nitrogen deficiency enhances expression of specific MYB and bHLH transcription factors and accumulation of end products in the flavonoid pathway. *Planta* 225, 1245–1253. doi: 10.1007/s00425-006-0414-x

Li, X., Chen, L., Hong, M., Zhang, Y., Zu, F., Wen, J., et al. (2012). A large insertion in bHLH transcription factor BrTT8 resulting in yellow seed coat in *Brassica rapa*. *PLoS ONE* 7:e44145. doi: 10.1371/journal.pone.0044145

Lila, M. A. (2004). Anthocyanins and human health: an *in vitro* investigative approach. *J. Biomed. Biotechnol.* 2004, 306–313. doi: 10.1155/S111072430440401X

Lindahl, M., Spetea, C., Hundal, T., Oppenheim, A. B., Adam, Z., and Andersson, B. (2000). The thylakoid FtsH protease plays a role in the light-induced turnover of the photosystem II D1 protein. *Plant Cell* 12, 419–431. doi: 10.1105/tpc.12.3.419

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5:3930. doi: 10.1038/ncomms4930

Manetas, Y. (2006). Why some leaves are anthocyanic and why most anthocyanic leaves are red? *Flora Morphol. Distribut. Funct. Ecol. Plants* 201, 163–177. doi: 10.1016/j.flora.2005.06.010

Martin, C., Butelli, E., Petroni, K., and Tonelli, C. (2011). How can research on plants contribute to promoting human health? *Plant Cell* 23, 1685–1699. doi: 10.1105/tpc.111.083279

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. *Plant J.* 55, 954–967. doi: 10.1111/j.1365-313x.2008.03565.x

Nagaharu, U. (1935). Genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jap. J. Bot.* 7, 389–452.

Nesi, N., Debeaujon, I., Jond, C., Pelletier, G., Caboche, M., and Lepiniec, L. (2000). The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in*Arabidopsis* siliques. *Plant Cell* 12, 1863–1878. doi: 10.1105/tpc.12.10.1863

Nicole, M. (2009). *The Photoprotective Role of Anthocyanin Pigmentin Leaf Tissues*. PhD Thesis, Wake Forest University.

Nixon, P. J., Michoux, F., Yu, J., Boehm, M., and Komenda, J. (2010). Recent advances in understanding the assembly and repair of photosystem II. *Ann. Bot.* 106, 1–16. doi: 10.1093/aob/mcq059

Parkin, I. A., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., et al. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol.* 15:R77. doi: 10.1186/gb-2014-15-6-r77

Payne, C. T., Zhang, F., and Lloyd, A. M. (2000). GL3 encodes a bHLH protein that regulates trichome development in Arabidopsis through interaction with GL1 and TTG1. *Genetics* 156, 1349–1362.

Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. *Plant Sci.* 181, 219–229. doi: 10.1016/j.plantsci.2011.05.009

Pietrini, F., Iannelli, M. A., and Massacci, A. (2002). Anthocyanin accumulation in the illuminated surface of maize leaves enhances protection from photo-inhibitory risks at low temperature, without further limitation to photosynthesis. *Plant Cell Environ.* 25, 1251–1259. doi: 10.1046/j.1365-3040.2002.00917.x

Ramsay, N. A., and Glover, B. J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 10, 63–70. doi: 10.1016/j.tplants.2004.12.011

Rubin, G., Tohge, T., Matsuda, F., Saito, K., and Scheible, W. R. (2009). Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in Arabidopsis. *Plant Cell* 21, 3567–3584. doi: 10.1105/tpc.109.067041

Sakamoto, W. (2003). Leaf-variegated mutations and their responsible genes in *Arabidopsis thaliana*. *Genes Genet. Syst.* 78, 1–9. doi: 10.1266/ggs.78.1

Steyn, W. J., Wand, S. J. E., Holcroft, D. M., and Jacobs, G. (2002). Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytol.* 155, 349–361. doi: 10.1046/j.1469-8137.2002.00482.x

Tohge, T., Nishiyama, Y., Hirai, M. Y., Yano, M., Nakajima, J., Awazuhara, M., et al. (2005). Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. *Plant J.* 42, 218–235. doi: 10.1111/j.1365-313X.2005.02371.x

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., et al. (1999). The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in Arabidopsis, encodes a WD40 repeat protein. *Plant Cell* 11, 1337–1350. doi: 10.1105/tpc.11.7.1337

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Xie, Q., Hu, Z., Zhang, Y., Tian, S., Wang, Z., Zhao, Z., et al. (2014). Accumulation and molecular regulation of anthocyanin in purple tumorous stem mustard (*Brassica juncea* var. *tumida Tsen et Lee). J. Agric. Food Chem.* 62, 7813–7821. doi: 10.1021/jf501790a

Xu, W., Dubos, C., and Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends Plant Sci.* 20, 176–185. doi: 10.1016/j.tplants.2014.12.001

Xu, W., Grain, D., Gourrierec, J., Harscoët, E., Berger, A., Jauvion, V., et al. (2013). Regulation of flavonoid biosynthesis involves an unexpected complex transcriptional regulation of *TT8* expression, in Arabidopsis. *New Phytol.* 198, 59–70. doi: 10.1111/nph.12142

Yu, F., Park, S., and Rodermel, S. R. (2004). The Arabidopsis FtsH metalloprotease gene family: interchangeability of subunits in chloroplast oligomeric complexes. *Plant J.* 37, 864–876. doi: 10.1111/j.1365-313X.2003.02014.x

Yu, F., Park, S., and Rodermel, S. R. (2005). Functional redundancy of AtFtsH metalloproteases in thylakoid membrane complexes. *Plant Physiol.* 138, 1957–1966. doi: 10.1104/pp.105.061234

Yuan, Y., Chiu, L. W., and Li, L. (2009). Transcriptional regulation of anthocyanin biosynthesis in red cabbage. *Planta* 230, 1141–1153. doi: 10.1007/s00425-009-1013-4

Zaltsman, A., Feder, A., and Adam, Z. (2005b). Developmental and light effects on the accumulation of FtsH protease in Arabidopsis chloroplasts—Implications for thylakoid formation and photosystem II maintenance. *Plant J.* 42, 609–617. doi: 10.1111/j.1365-313X.2005.02401.x

Zaltsman, A., Ori, N., and Adam, Z. (2005a). Two types of FtsH protease subunits are required for chloroplast biogenesis and photosystem II repair in Arabidopsis. *Plant Cell* 17, 2782–2790. doi: 10.1105/tpc.105.035071

Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., et al. (2015a). Disruption of a CAROTENOID CLEAVAGE DIOXYGENASE 4 gene converts flower colour from white to yellow in Brassica species. *New Phytol.* 206, 1513–1526. doi: 10.1111/nph.13335

Zhang, F., Gonzalez, A., Zhao, M., Payne, C. T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. *Development* 130, 4859–4869. doi: 10.1242/dev.00681

Zhang, Y., Butelli, E., De Stefano, R., Schoonbeek, H. J., Magusin, A., Pagliarani, C., et al. (2013). Anthocyanins double the shelf life of tomatoes by delaying over-ripening and reducing susceptibility to gray mold. *Curr. Biol.* 23, 1094–1100. doi: 10.1016/j.cub.2013.04.072

Zhang, Y., Chen, G., Dong, T., Pan, Y., Zhao, Z., Tian, S., et al. (2014). Anthocyanin accumulation and transcriptional regulation of anthocyanin biosynthesis in purple bok choy (*Brassica rapa* var chinensis). *J. Agric. Food Chem.* 62, 12366−12376. doi: 10.1021/jf503453e

Zhang, Y., de Stefano, R., Robine, M., Butelli, E., Bulling, K., Hill, L., et al. (2015b). Different ROS-scavenging properties of flavonoids determine their abilities to extend shelf life of tomato. *Plant Physiol.* 169, 1568–1583. doi: 10.1104/pp.15.00346

Zhang, Y., Hu, Z., Zhu, M., Zhu, Z., Wang, Z., Tian, S., et al. (2015c). Anthocyanin accumulation and molecular analysis of correlated genes in purple Kohlrabi (*Brassica oleracea* var. gongylodes L.). *J. Agric Food Chem.* 63, 4160–4169. doi: 10.1021/acs.jafc.5b00473

Zhou, Y., Guo, D., Li, J., Cheng, J., Zhou, H., Gu, C., et al. (2013). Coordinated regulation of anthocyanin biosynthesis through photorespiration and temperature in peach (*Prunus persica* f. *atropurpurea). Tree Genet. Genomes* 9, 265–278. doi: 10.1007/s11295-012-0552-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Mushtaq, Pan, Chen, Zhang, Ge and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 December 2015
doi: 10.3389/fpls.2015.01118

Chromosome Doubling of Microspore-Derived Plants from Cabbage (*Brassica oleracea* var. *capitata* L.) and Broccoli (*Brassica oleracea* var. *italica* L.)

*Suxia Yuan †, Yanbin Su †, Yumei Liu *, Zhansheng Li, Zhiyuan Fang, Limei Yang, Mu Zhuang, Yangyong Zhang, Honghao Lv and Peitian Sun*

Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China

OPEN ACCESS

Edited by:
Naser A. Anjum, University of Aveiro, Portugal

Reviewed by:
Margherita Irene Beruto, Istituto Regionale per la Floricoltura, Italy
Fernando Martinez, University of Seville, Spain

***Correspondence:**
Yumei Liu
liuyumei@caas.cn

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *09 September 2015*
Accepted: *25 November 2015*
Published: *22 December 2015*

Citation:
Yuan S, Su Y, Liu Y, Li Z, Fang Z, Yang L, Zhuang M, Zhang Y, Lv H and Sun P (2015) Chromosome Doubling of Microspore-Derived Plants from Cabbage (Brassica oleracea var. capitata L.) and Broccoli (Brassica oleracea var. italica L.). Front. Plant Sci. 6:1118. doi: 10.3389/fpls.2015.01118

Chromosome doubling of microspore-derived plants is an important factor in the practical application of microspore culture technology because breeding programs require a large number of genetically stable, homozygous doubled haploid plants with a high level of fertility. In the present paper, 29 populations of microspore-derived plantlets from cabbage (*Brassica oleracea* var. *capitata*) and broccoli (*Brassica oleracea* var. *italica*) were used to study the ploidy level and spontaneous chromosome doubling of these populations, the artificial chromosome doubling induced by colchicine, and the influence of tissue culture duration on the chromosomal ploidy of the microspore-derived regenerants. Spontaneous chromosome doubling occurred randomly and was genotype dependent. In the plant populations derived from microspores, there were haploids, diploids, and even a low frequency of polyploids and mixed-ploidy plantlets. The total spontaneous doubling in the 14 cabbage populations ranged from 0 to 76.9%, compared with 52.2 to 100% in the 15 broccoli populations. To improve the rate of chromosome doubling, an efficient and reliable artificial chromosome doubling protocol (i.e., the immersion of haploid plantlet roots in a colchicine solution) was developed for cabbage and broccoli microspore-derived haploids. The optimal chromosome doubling of the haploids was obtained with a solution of 0.2% colchicine for 9–12 h or 0.4% colchicine for 3–9 h for cabbage and 0.05% colchicine for 6–12 h for broccoli. This protocol produced chromosome doubling in over 50% of the haploid genotypes for most of the populations derived from cabbage and broccoli. Notably, after 1 or more years in tissue culture, the chromosomes of the haploids were doubled, and most of the haploids turned into doubled haploid or mixed-ploidy plants. This is the first report indicating that tissue culture duration can change the chromosomal ploidy of microspore-derived regenerants.

Keywords: cabbage (*Brassica oleracea* var. *capitata* L.), broccoli (*Brassica oleracea*. var. *italica* L.), microspore-derived plants, chromosome doubling, ploidy determination

INTRODUCTION

Microspore culture is an effective alternative technique for the production of doubled haploid (DH) parental lines to generate F_1 hybrids (Abercrombie et al., 2005). The development of DH lines accelerates the plant breeding process by saving time and labor (Ferrie and Caswell, 2011). In addition, DH lines can also be used for marker identification, gene mapping and various genetic manipulations (Forster et al., 2007; Ferrie and Möllers, 2011; Ferrie and Caswell, 2011). Consequently, this technique has been successfully used in cabbage and broccoli, and large-scale DH lines have been developed (Cao et al., 1990; Takahata and Keller, 1991; Duijs et al., 1992; Hansen, 1994; Pink, 1999; da Silva Dias, 2003; Yuan et al., 2009, 2011, 2012). Furthermore, some DH parental lines have even been introduced into breeding schemes (Hale et al., 2007; Lv et al., 2014). The procedure for DH production includes two major steps: haploid induction and chromosome doubling. Consequently, the chromosome doubling of haploids derived from microspores is an important step in the practical application of microspore culture technology. Microspore-derived haploids can spontaneously double their chromosomes during the very early stages of embryogenesis or can be induced to become DHs in the later stages of development (Palmer et al., 1996).

The ideal goal is to double the chromosome number of the original microspore and to then regenerate a plant from the resulting DH microspore. Theoretically, this would result in a stable, homozygous, completely fertile DH. Presently, the mechanism underlying this spontaneous chromosome doubling is unclear in many instances, with wide differences in the responses among and within species (Kasha, 2005). da Silva Dias (2003) and his group found 43–88% spontaneous diploids in broccoli and 7–91% spontaneous diploids in other coles. Because spontaneous doubling occurs randomly and is extremely genotype-dependent, it is important to ascertain the level of spontaneous diploids in the genotype used. da Silva Dias (2003) suggested that for genotypes with over 60% spontaneous doubling, it was not necessary to induce doubling. Obviously, for those genotypes with low spontaneous doubling rates, a successful chromosome-doubling process is essential for the production of homozygous plants after haploid plants are derived from the microspores of plants such as cabbage and broccoli. Various doubling agents have been studied, including caffeine (Thomas et al., 1997), nitrous oxide (Hansen et al., 1988), and the antimicrotubule herbicides trifluralin and amiprophos-methyl (APM; Hansen and Andersen, 1998b). However, the most commonly used chemical agent for chromosome doubling is colchicine (Niu et al., 2014), which disrupts mitosis by inhibiting the formation of spindle fibers and disturbing normal polar chromosomal migration, resulting in chromosome doubling (Jensen, 1974). da Silva Dias (2003) found that when the roots of plants derived from microspores were immersed in a 0.25% colchicine solution, a 53–71% doubling rate was achieved.

Previous studies examining the techniques of microspore culture in cole crops were focused strongly on improving embryogenesis and plant regeneration but ignored research on chromosome doubling. To date, successful microspore culture techniques have been established for cabbage and broccoli (Hansen, 2003; da Silva Dias, 2003; Yuan et al., 2009, 2011, 2012). Flower buds containing a large number of late uninucleate stage microspores and about 10–30% binucleate microspores were selected for microspore cultures in cabbage and broccoli (da Silva Dias, 2003). Heat shock pretreatment was shown to induce microspore embryogenesis in *Brassica oleracea* (Takahata and Keller, 1991; Duijs et al., 1992; da Silva Dias, 2001). In our previous research, the combination of cold pretreatment (4°C) for 1 or 2 days and heat shock (32.5°C) for 1 day significantly enhanced microspore embryogenesis in broccoli (Yuan et al., 2011), and 32.5°C for 1 or 2 days was optimal in cabbage (Yuan et al., 2012). NLN-13 medium and ½ NLN-13 medium were efficient on microspore embryogenesis in cabbage (Duijs et al., 1992) and broccoli (da Silva Dias, 2001), respectively. Furthermore, da Silva Dias (1999) reported that the addition of activated charcoal increased significantly embryo yields in nine genotypes of *B. oleracea*. Our previous study indicated that the combination of 10 mg l^{-1} gum arabic and 3 mM MES in NLN-13 at pH 6.4 significantly enhanced microspore embryogenesis in cabbage (Yuan et al., 2012). Based on the above-mentioned research, a large number of microspore-derived plant in cabbage and broccoli were obtained (Duijs et al., 1992; Yuan et al., 2009). For the useful application of microspore culture techniques in cabbage and broccoli breeding programs, the chromosome doubling of plants derived from microspores must be investigated and improved.

In the present paper, we used 29 populations of microspore-derived plantlets from cabbage (*Brassica oleracea* var. *capitata*) and broccoli (*Brassica oleracea* var. *italica*) to study the ploidy levels and spontaneous chromosome doubling in these populations. Additionally, we examined artificial chromosome doubling induced by colchicine and the influence of tissue culture duration on the chromosomal ploidy of the microspore-derived regenerants. Our objectives were to ascertain the spontaneous doubling of the populations, to assess the impact of tissue culture duration on chromosome doubling in the populations and to develop an efficient and reliable artificial chromosome doubling protocol for microspore-derived haploids of cabbage and broccoli.

MATERIALS AND METHODS

Plant Materials

From 2005 to 2013, the microspore culture of cabbage (*Brassica oleracea* var. *capitata*) and broccoli (*Brassica oleracea* var. *italica*) was undertaken from March to May at the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China. The microspore isolation and culture procedures were performed as previously described by Yuan et al. (2011, 2012). Briefly, NLN-13 (Yuan et al., 2012) and ½ NLN-13 (da Silva Dias, 2001) were used as liquid microspore culture media for cabbage and broccoli, respectively. The microspores were incubated in the dark at 32.5 ± 1°C for 1 day and then maintained at 25 ± 1°C in the dark. Cotyledonary embryos were obtained from 24-day-cultured microspores. The embryo culture and

plant regeneration were carried out according to the procedures described by Yuan et al. (2011). During subsequent tissue culture, the plantlets were subcultured using fresh solid MS-2 medium (Murashige and Skoog, 1962; 2% sucrose, 0.5% agar, 0.1 mg/l 1-naphthaleneacetic acid (NAA), 0.2 mg/l 6-benzylaminopurine (BAP), pH 5.8) every 2 months. A total of 29 populations of cabbage and broccoli microspore-derived plantlets were obtained (**Table 1**). One copy of the microspore-derived plants was continually grown in a tissue culture room, while another copy was transferred to a greenhouse after October and then planted in a field to allow flowering and set seed.

Chromosome Doubling

Spontaneous Doubling

To accurately determine the chromosomal ploidy of the regenerants, the ploidy level of each population mentioned above was identified using two methods: chloroplast counting and morphological identification. Chloroplast counting was used for early ploidy identification in tissue culture, and then to divide each population into haploid group, haploid group, polyploid group and ploid-mixed group in order to facilitate morphological identification in flowering.

Artificial Chromosome Doubling

The haploid genotype 04M1-93 derived from cabbage "Zhonggan No. 11" and the haploid genotype 05B743-49 derived from broccoli "TI-111" were tested in this experiment. To obtain sufficient plants for artificial chromosome doubling research, the two haploids were propagated in MS medium (Murashige and Skoog, 1962; 3% sucrose, 0.5% agar, 0.1 mg/l NAA, 1 mg/l BAP, pH 5.8). In the mid of September, 2007, all of the haploid plants were cut free from the hypocotyl tissue and were transferred to glass growth vessels containing solid ½ MS-2 medium (Murashige and Skoog, 1962; the concentration of the major salts was reduced to 50% compared with MS medium, 2% sucrose, 0.5% agar, 0.1 mg/l NAA, 0.1 mg/l indole-3-butytric acid (IBA), and pH 5.8), in which rooting took place.

The colchicine treatment was as follows: 2 weeks after rooting, the rooted plants were removed from the medium and washed completely with warm tap water; the roots were then trimmed to a length of 1–2 cm and immersed in a working solution of colchicine (supplemented with 2% DMSO). The plants were placed under intense light at 25°C. Next, the solution was poured off, and the roots were rinsed thoroughly in tap water. The treated plants were replanted in a soil mixture (peat soil: perlite: vermiculite, 8:1:1) in a pot and maintained for approximately 2 more weeks in a room at 24°C with a 16-h photoperiod, with low light intensity and high humidity. This was followed by gradual adaptation to greenhouse conditions. Approximately 6 weeks later, the plants were transferred to a cold frame for the duration of the winter. In April of the next year, the plants started flowering, and the ploidy level of the plants was determined using morphological identification.

In this experiment, four concentration levels of colchicine solution were tested: 0.05, 0.1, 0.2, and 0.4%. The plants were treated for 3, 6, 9, and 12 h, in each colchicine solution. Fifteen plants of every haploid genotype were used in each treatment. The same number of plants of each haploid genotype grown without colchicine served as the controls.

Evaluation of Artificial Chromosome Doubling

Based on the experimental results of 2.2.2, artificial chromosome doubling induced by colchicine was applied to 576 haploid genotypes derived from six cabbage and six broccoli populations to test the chromosome doubling efficiency. In this experiment, 4–8 plants of each haploid genotype, i.e., a total of 3128 haploid plants, were treated with a colchicine solution.

TABLE 1 | Microspore-derived cabbage and broccoli plantlet populations with different genotypes were obtained from 2005 to 2013.

Population genotype		Year	Population genotype		Year
Cabbage	Zhonggan No. 11	2005∼2007, 2009	Broccoli	Lüxiu	2005∼2007
	8398	2005∼2007		Lüxing	2005∼2007
	Xiwang	2005∼2007		TI-111	2005∼2007
	Zhonggan No. 18	2005∼2007		TI-089	2005∼2007
	08F8	2009		Lüyu	2005∼2007
	08F9	2009		Yuguan	2005∼2007
	08F15	2009		09B411	2010
	08F16	2009		09B524	2010
	08F17	2009		10B635	2011
	08F18	2009		10B644	2011
	08F24	2009		10B665	2011
	08F25	2009		10B908	2011
	Gen 165	2012		10B1092	2011
				12B520	2013
				12B1153	2013

Determination of the Ploidy Level

Chloroplast Number

In the tissue cultures, chloroplast counting (Yuan et al., 2009) was used to determine the initial ploidy level of the populations. In cole crops, haploid plants have at most 10 chloroplasts, diploid plants have 11–15, and polyploid plants have more than 15 chloroplasts.

Morphological Identification

In the mid of September each year, another copy of each of the above-mentioned microspore-derived plantlets obtained in the same year was cut free from the hypocotyl tissue and transferred to a glass growth vessel containing solid ½ MS-2 medium, in which rooting took place. After 2 weeks, the rooted plantlets were transferred to a soil-perlite mixture in a pot. This transfer was then followed by the gradual adaptation to greenhouse conditions. At the end of January of the following year, the plantlets were transferred to a cold frame for the winter. In April of that year, the regenerated plants began flowering, and the population ploidy level was determined using morphological identification.

The sizes of the plants, buds and flowers were observed in the field from 2006 to 2014. The presence or absence of pollen was an additional morphological feature that was determined. Normal diploid cabbage and broccoli plants were used as controls. The characteristics of the regenerated plants with different ploidy levels were described according to the following (**Figure 1**):

Haploid: The growth potential of the plant is weaker, and the plant size is smaller. The flower buds are smaller, flatter and without pollen. The stamens are missing, or the stamen development is not normal.
Diploid: The plant grows normally and has pollen. The stamens and pistils are normal.
Triploid: The flower buds are smaller, flatter and without pollen. The stamens are missing, or the stamen development is not normal.
Tetraploid: The plant has pollen and stronger growth vigor. The plant size and flower buds are larger. The stamens and pistil are normal.



FIGURE 1 | The characteristics of the inflorescences, buds and flowers from different chromosome ploidy plants derived from cabbage "Zhonggan No. 11" microspores. (A) Flowers from different chromosome ploidy plants; **(B)** Buds from different chromosome ploidy plants; **(C)** Inflorescences from different chromosome ploidy plants; (a) Tetraploid; (b) Triploid; (c) Diploid; (d) Haploid.

Data Analysis

The data were analyzed using Microsoft Excel 2003.

RESULTS

Ploidy Level and Spontaneous Chromosome Doubling of the Populations

In total, 1717 regenerants of cabbage and 622 of broccoli were derived and investigated for their ploidy level (**Table 2**). Spontaneous doubling occurred randomly and was extremely genotype dependent. In the microspore-derived populations, in addition to haploids and diploids, there was a low frequency of polyploids and mixed-ploidy plantlets (a plant having both haploid and diploid branches simultaneously; **Figure 2**). For the 14 cabbage genotypes, the spontaneous diploid rates for the populations were in the range of 0–76.9%, and the total spontaneous doubling (including diploids, polyploids, and mixed-ploidy plantlets) ranged from 0 to 84.6%. By contrast, in the 15 broccoli populations, the spontaneous diploid rate was 50.6–100%, and the total spontaneous doubling ranged from 52.2 to 100%. In this experiment, cabbage exhibited a larger variation in the rate of spontaneous doubling in the microspore-derived regenerated populations than did broccoli.

Artificial Chromosome Doubling Induced by Colchicine

The data presented in **Table 3** indicate that the plant survival rate gradually decreased with an increase in the concentration of the colchicine solution. Similarly, for the same concentration of colchicine, the plant survival rate decreased with an increase of the treatment time. Therefore, a higher concentration of colchicine or a longer duration of colchicine treatment had negative effects on the survival of the haploids; this phenomenon was more obvious in broccoli than in cabbage. In addition, colchicine not only induced the doubling of haploids but also induced a certain frequency of haploids to become mixed-ploidy plants (**Figure 2**).

The 0.2% colchicine solution produced negative effects on the survival of the 04M1-93 haploids derived from "Zhonggan No. 11" cabbage; however, the 0.05% colchicine solution produced negative effects on the survival of the 05B743-49 haploids derived from broccoli "TI-111."

When the cabbage haploids were treated with the 0.2% colchicine solution for 9 h, the DH frequency was the highest, up to 50%, followed by the 0.4% colchicine solution for 3 h. However, when the haploids were treated with the 0.4% colchicine solution for 6 h, the total chromosome doubling (including diploids and mixed-ploidy plantlets) was the highest, followed by the 0.4% colchicine for 9 h. Mixed-ploidy plants have diploid branches, the flowers of which exhibit normal fertility and can be pollinated; consequently, these plants have value for breeding. Considering the DH frequency, survival rate and total chromosome doubling, we concluded that a better chromosome doubling effect was obtained when the haploid plants were treated with 0.2% colchicine for 9–12 h or 0.4% colchicine for 3–9 h. The DH

TABLE 2 | Chromosomal ploidy levels of cabbage and broccoli populations derived from microspores.

Materials		Number of populations	Haploid		Diploid		Polyploid		Mixed-ploidy		Rate of chromosome doubling (%)
			Number of plants	Frequency (%)	Number of plants	Frequency (%)	Number of plants	Frequency (%)	Number of plants	Frequency (%)	
Cabbage	Zhonggan No. 11	210	35	16.7	158	75.2	2	1.0	15	7.1	83.3
	8398	381	205	53.8	125	32.8	5	1.3	46	12.1	46.2
	Xiwang	52	8	15.4	40	76.9	0	0	4	7.7	84.6
	Zhonggan No. 8	53	11	20.8	38	71.7	1	1.9	3	5.6	79.2
	Zhonggan No. 21	34	22	64.7	10	29.4	2	5.9	0	0.0	35.3
	08F8	106	52	49.1	49	46.2	2	1.9	3	2.8	50.9
	08F15	3	3	100.0	0	0.0	0	0.0	0	0.0	0.0
	08F16	2	2	100.0	0	0.0	0	0.0	0	0.0	0.0
	08F17	3	2	66.7	1	33.3	0	0.0	0	0.0	33.3
	08F9	116	95	81.9	12	10.3	4	3.5	5	4.3	18.1
	08F18	99	31	31.3	62	62.6	4	4.0	2	2.0	68.7
	08F24	380	171	45.0	192	50.5	3	0.8	14	3.7	55.0
	08F25	266	123	46.2	124	46.6	11	4.2	8	3.0	53.8
	gen 165	12	12	100.0	0	0.0	0	0.0	0	0.0	0.0
Broccoli	Lüxiu	23	11	47.8	12	52.2	0	0.0	0	0.0	52.2
	Lüxing	163	8	4.9	150	92.0	1	0.6	4	2.5	95.1
	TI-111	91	42	46.1	46	50.6	0	0.0	3	3.3	53.9
	TI-089	20	0	0.0	20	100.0	0	0.0	0	0.0	100.0
	Lüyu	152	4	2.6	144	94.7	1	0.7	3	2.0	97.4
	Yuguan	3	0	0.0	3	100.0	0	0	0	0.0	100.0
	09B411	32	6	18.8	23	71.9	1	3.1	2	6.2	81.2
	09B524	4	0	0.0	4	100.0	0	0.0	0	0.0	100.0
	10B908	36	16	44.4	19	52.8	0	0.0	1	2.8	55.6
	10B1092	36	14	38.9	21	58.3	0	0.0	1	2.8	61.1
	10B635	24	5	20.8	19	79.2	0	0.0	0	0.0	79.2
	10B644	21	5	23.8	15	71.4	1	4.8	0	0.0	76.2
	10B665	2	0	0.0	2	100.0	0	0.0	0	0.0	100.0
	12B520	12	3	25.0	9	75.0	0	0.0	0	0.0	75.0
	12B1153	3	0	0.0	3	100.0	0	0.0	0	0.0	100.0

frequency ranged from 23.1 to 50.0%, the total chromosome doubling ranged from 58.3 to 84.6%, and the survival rate ranged from 73.3 to 93.3%.

Similarly, for the 05B743-49 haploids derived from broccoli "TI-111," we concluded that a better chromosome doubling effect was obtained when the haploid plants were treated with 0.05% colchicine for 6–12 h; the DH frequency ranged from 16.7 to 37.5%, the total chromosome doubling ranged from 54.5 to 75.0%, and the survival rate ranged from 57.1 to 80.0%.

Evaluation of Chromosome Doubling in Cabbage and Broccoli Induced by Colchicine

According to the experimental results mentioned above, a 0.2% colchicine treatment for 9–12 h was used for cabbage, and a 0.05% colchicine treatment for 9–12 h was used for broccoli. The plant survival rate of the cabbage population "08F9" was low, at 61.9%, but the survival rates for the 11 other populations ranged from 77.3 to 87.0% (**Table 4**). Among the surviving plants, with the exception of the cabbage population "08F9" and the broccoli population "09B411," which had low total chromosome doubling rates of only 24.6 and 31.7%, respectively, the five other cabbage populations had total chromosome doubling rates ranging from 42.3 to 57.7%, and those for the five other broccoli populations ranged from 43.8 to 60.0%. It is worth mentioning that a higher frequency of chromosome-doubled genotypes was obtained for all of the haploid genotypes of each population compared with the total chromosome doubling efficiency of the surviving plants. For the cabbage population "08F9" and the broccoli population "09B411," the frequencies of the chromosome-doubled genotypes were 40.9 and 39.3%, respectively. For the five other cabbage and five other broccoli populations, the frequencies of the chromosome-doubled genotypes ranged from 49.2 to 69.6% and 64.3 to 72.7%, respectively.



FIGURE 2 | The morphological characteristics of mixed-ploidy plants (a plant having both haploid and diploid branches simultaneously) derived from cabbage "Zhonggan No. 11" microspores. (A) Flowers from a haploid branch; **(B)** Flowers from a diploid branch.

Effect of Tissue Culture Duration on the Chromosome Doubling of the Population

An experiment was performed to determine whether the tissue culture duration could change the chromosomal ploidy of the microspore-derived regenerants.

In this experiment, 9 populations of cabbage and 3 populations of broccoli were tested, and the ploidy variation of the plantlets was observed during the tissue culture. Notably, the chromosome doubling rate of the population gradually increased with an increase in the tissue culture time (**Table 5**). After 1 or more years of tissue culture, the number of chromosomes of the haploids doubled, and most of the haploids turned into DHs or mixed-ploidy plants (data not shown).

DISCUSSION

The chromosome doubling of microspore-derived haploids is an important factor in the practical application of microspore culture technology because breeding programs require a large number of genetically stable, homozygous DH plants with a high level of fertility. Theoretically, a microspore only carries half of the chromosomes of the somatic donor plant, and a plantlet derived from a microspore should be haploid. However, in practice, microspore-derived populations, in addition to haploids, also contain a certain proportion of diploids and even a small number of polyploids. Furthermore, in our experiments, in addition to haploids, diploids and polyploids, a low frequency of mixed-ploidy plantlets (a plant having both haploid and diploid branches simultaneously) was also obtained in cabbage and broccoli. The spontaneous doubling rate varies among crop species and genotypes. In *Brassica napus,* the spontaneous chromosome doubling rates of the microspores range from 10 to 26% (Chen et al., 1994), whereas higher spontaneous doubling rates (50–70%) were observed in Chinese cabbage (Zhang and Takahata, 2001). In broccoli, the diploid rate ranges from 43 to 88%, and in other coles, this rate ranges from 7 to 91% (da Silva Dias, 2003). In our research, the spontaneous diploid rate in cabbage ranged from 0 to 76.9%, and the total spontaneous doubling ranged from 0 to 84.6%. However, in broccoli, the spontaneous diploid rate ranged from 50.6 to 100%, and the total spontaneous doubling ranged from 52.2 to 100% (**Table 2**). Obviously, there was a larger range in the spontaneous chromosome doubling in the microspore-derived regenerated cabbage populations compared with those of broccoli. This is in good agreement with the results presented by da Silva Dias (2003).

A high proportion of spontaneous DH plants is particularly beneficial in breeding because there is no need to use colchicine to double the haploid chromosomes. Spontaneous doubling saves time and labor. To date, the mechanism of spontaneous chromosome doubling remains unclear, while in the mitotic divisions of microspores, the induction of androgenesis and chromosome doubling both appear to involve changes in microfilaments and microtubules. Microfilaments and microtubules may be responsible for the nuclear migration around the uni-nucleate microspore wall. If a pretreatment system for inducing embryogenesis disrupts the microtubules, this type of treatment might lead to chromosome doubling (Kasha, 2005). Colchicine and a number of other anti-microtubule agents have been used to improve chromosome doubling in microspore cultures of *Brassica* crops. Zhao et al. (1996) and Zhou et al. (2002) found that treatment with colchicine instead of heat shock improved the diploid frequency of microspore-derived *Brassica napus* plants. Li and Devaux (2003) found that cold and mannitol pretreatments of barley microspore cultures resulted in high rates of DHs, regardless of the explant (anther or spike) used. Kasha et al. (2001) concluded that the nuclear fusion of microspores was the main mechanism of spontaneous chromosome doubling in barley isolated microspore cultures following mannitol and cold pretreatments. In our previous study, the combination of cold pretreatment and heat shock resulted in a population with more spontaneous DHs compared with heat shock alone (Yuan et al., 2011).

Furthermore, the microspore culture stage could also influence spontaneous chromosome doubling. For example, culturing of early uni-nucleate stages produced predominantly haploid progeny, whereas culturing of bi-nucleate stages produced more DHs and polyploids (Kasha, 2005). Soriano et al. (2007) found that the spontaneous doubling rates (54–66%) of wheat microspores during the late uni-nucleate to early bi-nucleate stages were higher than those (33%) in the mid to late uni-nucleate stages. This may be related to the late uni-nucleate to early bi-nucleate stages being the best microspore stages for the highest embryo induction (Pechan and Keller, 1988).

Although, several factors can affect spontaneous chromosome doubling in microspores, the same microspore culture stage and type of pretreatment were used for the genotypes of the cabbage and broccoli microspore cultures in our study. Our

TABLE 3 | Efficiency of chromosome doubling of different haploids induced by colchicine.

Materials		Concentration of colchicine (%)	Treatment duration (h)	Number of plants	Number of surviving plants	Chromosomal ploidy level			Frequency of chromosome doubling (%)		Survival rate (%)
						Haploid	Diploid	Mixed-ploidy	Diploid frequency	Total chromosome doubling	
Cabbage	04M1-93	0 (CK)	0	15	15	15	0	0	0.0	0.0	100.0
		0.05	3	15	15	12	0	3	0.0	20.0	100.0
			6	15	15	12	0	3	0.0	20.0	100.0
			9	15	15	7	0	8	0.0	53.3	100.0
			12	15	15	7	1	7	6.7	53.3	100.0
		0.10	3	14	14	10	0	4	0.0	28.6	100.0
			6	15	15	9	0	6	0.0	40.0	100.0
			9	14	14	4	3	7	21.4	71.4	100.0
			12	15	15	8	4	3	26.7	46.7	100.0
		0.20	3	15	15	5	0	10	0.0	66.7	100.0
			6	15	13	7	3	3	23.1	46.2	86.7
			9	15	12	5	6	1	50.0	58.3	80.0
			12	15	12	3	3	6	25.0	75.0	80.0
		0.40	3	15	14	3	5	6	35.7	78.6	93.3
			6	15	13	2	3	8	23.1	84.6	86.7
			9	15	11	2	3	6	27.3	81.8	73.3
			12	15	6	3	0	3	0.0	50.0	40.0
Broccoli	05B743-49	0 (CK)	0	4	4	4	0	0	0.0	0.0	100.0
		0.05	3	15	14	8	4	2	28.6	42.9	93.3
			6	15	11	5	3	3	27.3	54.5	73.3
			9	14	8	2	3	3	37.5	75.0	57.1
			12	15	12	5	2	5	16.7	58.3	80.0
		0.10	3	15	13	7	1	5	7.7	46.2	86.7
			6	15	6	1	1	4	16.7	83.3	40.0
			9	15	6	1	1	4	16.7	83.3	40.0
			12	15	3	0	2	1	66.7	100.0	20.0
		0.20	3	15	7	2	1	3	14.3	57.1	46.7
			6	15	6	2	3	1	50.0	66.7	40.0
			9	15	2	0	0	1	0.0	50.0	13.3
			12	15	2	0	1	1	50.0	100.0	13.3
		0.40	3	15	3	1	1	1	33.3	66.7	20.0
			6	15	1	0	0	1	0.0	100.0	6.7
			9	15	0	0	0	0	0.0	0.0	0.0
			12	15	0	0	0	0	0.0	0.0	0.0

results indicate that the spontaneous chromosome doubling rate varies among *B. oleracea* crops and genotypes.

da Silva Dias (2003) suggested that is not necessary to induce doubling in genotypes with over 60% spontaneous doubling. In our research, the spontaneous doubling rate of 10 of the 14 microspore-derived populations (71.4%) of cabbage was less than 60% (**Table 2**). It is clear that the frequency of spontaneous doubling can be very low in cabbage. The spontaneous doubling rate exceeded 60% in most of the broccoli populations in the experiment (**Table 2**). However, a successful chromosome-doubling process is necessary for genotypes with low spontaneous doubling. Therefore, the development of efficient chromosome-doubling protocols is essential for the useful application of DH plants in cabbage and broccoli breeding programs.

Chromosome doubling can be induced in the early stages of gametic embryogenesis or during the developmental stage of haploid plantlets (Ferrie, 2003). Rudolf et al. (1999) found that *in vitro* treatment with anti-microtubule agents can enhance chromosome doubling, but this method can also reduce embryogenesis (Ferrie, 2003) and regenerant frequency (Hansen and Andersen, 1998a) in microspore cultures. Furthermore, this approach can increase the contamination of microspore cultures. In our study, haploid plantlet roots were immersed in a colchicine solution. This method is simple, convenient and targeted. This procedure produced double chromosomes in over 50% of the

TABLE 4 | Application of chromosome doubling in different cabbage and broccoli genotypes induced by colchicine.

Materials		Number of haploid genotypes	Number of plants	Number of surviving plants	Chromosomal ploidy level			Frequency of chromosome doubling (%)		Plant survival rate (%)	Total chromosome chromosome doubling	
					Haploid	Diploid	Mixed-ploidy	Diploid frequency	Total chromosome doubling		Number of genotypes	Frequency of genotypes (%)
Cabbage	08F8	72	360	288	146	124	18	43.1	49.3	80.0	39	54.2
	08F9	44	210	130	98	28	4	21.5	24.6	61.9	18	40.9
	08F18	44	189	156	66	84	6	53.8	57.7	82.5	30	68.2
	08F24	129	821	660	330	318	12	48.2	50.0	80.4	76	58.9
	08F25	118	682	532	307	211	14	39.7	42.3	78.0	58	49.2
	Zhonggan No. 11	23	92	80	44	40	6	50.0	57.5	87.0	16	69.6
Broccoli	09B411	56	286	243	166	55	22	22.6	31.7	85.0	22	39.3
	10B635	22	122	101	55	32	14	31.7	45.5	82.8	16	72.7
	10B644	17	88	68	36	24	8	35.3	47.1	77.3	11	64.7
	10B908	20	108	89	50	32	7	36.0	43.8	82.4	13	65.0
	10B1092	28	146	114	54	48	12	42.1	52.6	78.1	18	64.3
	12B520	3	24	20	8	11	1	55.0	60.0	83.3	2	66.7

TABLE 5 | Spontaneous doubling of populations during different tissue culture years.

Materials		Number of populations	Year			
			2010	2011	2012	2013
Cabbage	08F8	106	50.9	66.2	74.3	82.4
	08F9	116	18.1	24.0	38.5	47.8
	08F15	3	0.0	50.0	50.0	100.0
	08F16	2	0.0	25.0	50.0	100.0
	08F17	3	33.3	33.3	67.7	67.7
	08F18	99	68.7	78.6	82.4	96.0
	08F24	380	55.0	76.6	84.4	92.2
	08F25	266	53.8	75.0	82.4	90.0
	Zhonggan 11	145	65.5	77.4	85.2	100.0
Broccoli	09B411	32	–	81.3	93.6	100.0
	10B908	36	–	–	55.6	77.8
	10B1092	36	–	–	61.1	88.9

haploid genotypes for most of the populations (75%) derived from cabbage and broccoli (**Table 4**).

Many factors affect the chromosome-doubling process, such as the colchicine concentration, treatment duration, the addition of other synthetic compounds, and the developmental stage of the plants.

A higher colchicine concentration can increase the doubling rate, but a higher concentration also results in a low survival rate and increased cost. Similarly, the duration of colchicine exposure can have a considerable effect on the induction of chromosome doubling and the survival rate. The optimal colchicine treatment should result in a high plant survival rate and a high rate of chromosome doubling. In this study, 0.2% colchicine for 9–12 h or 0.4% colchicine for 3–9 h for cabbage and 0.05% colchicine for 6–12 h for broccoli produced optimal chromosome haploid doubling effects (**Table 3**).

In our study, the immersion of haploid plantlet roots in a colchicine solution was used to induce chromosome doubling; therefore, the haploid plants used for this purpose must have good root systems. The development of the root systems can affect the colchicine solution absorption efficiency, chromosome doubling and survival rates. Prior to the colchicine treatment, we selected haploid plants with good root systems and then trimmed the roots to better absorb the colchicine solution. Furthermore, 2% DMSO was added to the colchicine solution to increase the root absorption rate.

It is well known that two approaches can be used for chromosome doubling, i.e., spontaneous doubling and artificial chromosome doubling induced by colchicine or other anti-microtubule agents. In our study, which spanned a period of 4 years, we noted that after 1 or more years of tissue culture, the chromosome content of the haploids was doubled, and most of the haploid plants became DHs or mixed-ploidy plants. This phenomenon indicates that the chromosome number of haploids derived from cabbage and broccoli microspores is not stable and can easily be induced to change as a result of external conditions.

This is the first report suggesting that tissue culture duration can change the chromosomal ploidy of the microspore-derived regenerants.

AUTHOR CONTRIBUTIONS

SY performed the experiments, analyzed the data, wrote and revised the manuscript; YS performed the experiments, analyzed the data and revised the manuscript; YL designed the research and critically edited the manuscript; and ZL, ZF, LY, MZ, YZ, HL, and PS planted and managed the plants. All authors approved the final manuscript.

ACKNOWLEDGMENTS

This work was funded by the Chinese National Natural Science Foundation (30871708, 31372067), the Earmarked Fund for Modern Agro-industry Technology Research System (CARS-25-A), the Chinese National Key Technology R&D Program (2013BAD01B04), the National High Technology Research and Development Program (863 Program; 2012AA100104 and 2012AA100105) and the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, P. R. China. We also acknowledge partial funding from the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS).

REFERENCES

Abercrombie, J. M., Farnham, M. W., and Rushing, J. W. (2005). Genetic combining ability of glucoraphanin level and other horticultural traits of broccoli. *Euphytica* 143, 145–151. doi: 10.1007/s10681-005-3059-0

Cao, M. Q., Charlot, F., and Doré, C. (1990). Embryogenesis and plant regeneration in sauerkraut cabbage (*Brassica oleracea* L. subsp. *Capitata*) by *in vitro* culture of isolated microspores. *C. R. Acad. Sci. Paris Ser.* 310, 203–209.

Chen, Z. Z., Snyder, S., Fan, Z. G., and Loh, W. H. (1994). Efficient production of doubled haploid plants through chromosome doubling of isolated microspores in *Brassica napus*. *Plant Breed.* 113, 217–221. doi: 10.1111/j.1439-0523.1994.tb00725.x

da Silva Dias, J. C. (1999). Effect of activated charcoal on *Brassica oleracea* microspore culture embryogenesis. *Euphytica* 108, 65–69.

da Silva Dias, J. C. (2001). Effect of incubation temperature regimes and culture medium on broccoli microspore culture embryogenesis. *Euphytica* 119, 389–394. doi: 10.1023/A:1017563915319

da Silva Dias, J. C. (2003). "Protocol for broccoli microspore culture," in *Doubled Haploid Production in Crop Plants*, eds M. Maluszynski, K. J. Kasha, B. P. Forster, and I. Szarejko (Dordrecht: Kluwer Academic Publishers), 195–204.

Duijs, J. G., Voorrips, R. E., Visser, D. L., and Custers, J. B. M. (1992). Microspore culture is successful in most crop types of *Brassica oleracea* L. *Euphytica* 60, 45–55.

Ferrie, A. (2003). "Microspore culture of Brassica species," in *Doubled Haploid Production in Crop Plants,* eds M. Maluszynski, K. J. Kasha, B. P. Forster, and I. Szarejko (Dordrecht: Kluwer Academic Publishers), 205–215.

Ferrie, A. M. R., and Caswell, K. L. (2011). Isolated microspore culture techniques and recent progress for haploid and doubled haploid plant production. *Plant Cell Tiss. Organ Cult.* 104, 301–309. doi: 10.1007/s11240-010-9800-y

Ferrie, A. M. R., and Möllers, C. (2011). Haploids and doubled haploids in Brassica spp. for genetic and genomic research. *Plant Cell Tiss. Organ Cult.* 104, 375–386. doi: 10.1007/s11240-010-9831-4

Forster, B. P., Heberle-Bors, E., Kasha, K. J., and Touraev, A. (2007). The resurgence of haploids in higher plants. *Trends Plant Sci.* 12, 368–375. doi: 10.1016/j.tplants.2007.06.007

Hale, A. L., Farnham, M. W., Nzaramba, M. N., and Kimbeng, C. A. (2007). Heterosis for horticultural traits in broccoli. *Theor. Appl. Genet.* 115, 351–360. doi: 10.1007/s00122-007-0569-2

Hansen, F. L., Andersen, S. B., Due, I. K., and Olesen, A. (1988). Nitrous oxide as a possible alternative agent for chromosome doubling of wheat haploids. *Plant Sci.* 54, 219–222. doi: 10.1016/0168-9452(88)90116-1

Hansen, M. (1994). "Gametic embryogenesis in *Brassica*: optimization of production and germination of embryos," in *Proceedings of the Int Colloquium on Impact of Plant Biotechnology on Agriculture,* eds B. Javornik, B. Bohanec, I. Kreft (Rogla: Centre for Plant Biotechnology and Breeding; University of Ljubjana), 15–18.

Hansen, M. (2003). "Protocol for microspore culture in Brassica," in *Doubled Haploid Production in Crop Plants*, eds M. Maluszynski, K. J. Kasha, B. P. Forster, and I. Szarejko (Dordrecht: Kluwer Academic Publishers), 195–204.

Hansen, N. J. P., and Andersen, S. B. (1998a). *In vitro* chromosome doubling with colchicine during microspore culture in wheat (*Triticum aestivum* L.). *Euphytica* 102, 101–108. doi: 10.1023/A:1018348816205

Hansen, N. J. P., and Andersen, S. B. (1998b). Efficient production of doubled haploid wheat plants by *in vitro* treatment of microspores with trifluralin or APM. *Plant Breed.* 117, 401–405. doi: 10.1111/j.1439-0523.1998.tb01963.x

Jensen, C. J. (1974). "Chromosome doubling techniques in haploids," in *Haploids in Higher Plants Advances and Potential,* ed K. L. Kasha (Guelph, ON: University of Guelph), 153–190.

Kasha, K. J. (2005). "Chromosome doubling and recovery of doubled haploid plants," in *Haploids in Crop Improvement?. Biotechnology in Agriculture and Forestry, Vol. 56,* eds C. E. Palmer, W. A. Keller, K. J. Kasha (Berlin: Springer), 123–152.

Kasha, K. J., Hu, T. C., Oro, R., Simion, E., and Shim, Y. S. (2001). Nuclear fusion leads to chromosome doubling during mannitol pretreatment of barley (*Hordeum vulgare* L.) microspores. *J. Exp. Bot.* 52, 1227–1238. doi: 10.1093/jexbot/52.359.1227

Li, H. C., and Devaux, P. (2003). High frequency regeneration of barley doubled haploid plants from isolated microspore culture. *Plant Sci.* 164, 379–386. doi: 10.1016/S0168-9452(02)00424-7

Lv, H., Wang, Q., Zhang, Y., Yang, L., Fang, Z., Wang, X., et al. (2014). Linkage map construction using InDel and SSR markers and QTL analysis of heading traits in *Brassica oleracea* var. *capitata* L. *Mol. Breed.* 34, 87–98. doi: 10.1007/s11032-014-0019-1

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Plant Physiol.* 15, 473–497.

Niu, Z., Jiang, A., Abu Hammad, W., Oladzadabbasabadi, A., Xu, S. S., Mergoum, M., et al. (2014). Review of doubled haploid production in durum and common wheat through wheat × maize hybridization. *Plant Breed.* 133, 313–320. doi: 10.1111/pbr.12162

Palmer, C. E., Keller, W. A., and Arnison, P. G. (1996). "Utilization of Brassica haploids," in *In Vitro Haploid Production in Higher Plants*, eds S. M. Jain, S. K. Sopory, and R. E. Veilleux (Dordrecht: Kluwer Academic Publishers), 173–192.

Pechan, P. M., and Keller, W. A. (1988). Identification of potentially embryogenic microspores in *Brassica napus*. *Physiol. Plant* 74, 377–384.

Pink, D. (1999). "Application of doubled haploid technology and DNA Markers in breeding for Clubroot resistance in *Brassica oleracea*," in *COST-824 Gametic Embryogenesis Workshop, Book of Abstracts*, ed M. Hansen (Krakow: European Commission), 5–7.

Rudolf, K., Bohanec, B., and Hansen, M. (1999). Microspore culture of white cabbage, Brassica oleracea var. *capitata* L.: genetic improvement of non-responsive cultivars and effect of genome doubling agents. *Plant Breed.* 118, 237–241.

Soriano, M., Cistué, L., Vallés, M. P., and Castillo, A. M. (2007). Effects of colchicine on anther and microspore culture of bread wheat (*Triticum aestivum* L.). *Plant Cell Tiss. Organ Cult.* 91, 225–234. doi: 10.1007/s11240-007-9288-2

Takahata, Y., and Keller, W. A. (1991). High frequency embryogenesis and plant regeneration in isolated microspore culture of *Brassica oleracea* L. *Plant Sci.* 74, 235–242. doi: 10.1016/0168-9452(91)90051-9

Thomas, J., Chen, Q., and Howes, N. (1997). Chromosome doubling of haploids of common wheat with caffeine. *Genome* 40, 552–558. doi: 10.1139/g97-072

Yuan, S., Liu, Y., Fang, Z., Yang, L., Zhuang, M., Zhang, Y., et al. (2009). Study on the relationship between the ploidy level of microspore-derived plants and the number of chloroplast in stomatal guard cells in *Brassica oleracea*. *Agric. Sci. China* 8, 939–946. doi: 10.1016/S1671-2927(08)60298-9

Yuan, S., Su, Y., Liu, Y., Fang, Z., Yang, L., Zhuang, M., et al. (2012). Effects of pH, MES, arabinogalactan-proteins on microspore cultures in white cabbage. *Plant Cell Tiss. Organ Cult.* 110, 69–76. doi: 10.1007/s11240-012-0131-z

Yuan, S. X., Liu, Y. M., Fang, Z. Y., Yang, L. M., Zhuang, M., Zhang, Y. Y., et al. (2011). Effect of combined cold pretreatment and heat shock on microspore cultures in broccoli. *Plant Breed.* 130, 80–85. doi: 10.1111/j.1439-0523.2009.01754.x

Zhang, F. L., and Takahata, Y. (2001). Inheritance of microspore embryogenic ability in *Brassica* crops. *Theor. Appl. Genet.* 103, 254–258. doi: 10.1007/s001220100602

Zhao, J., Simmonds, D. H., and Newcomb, W. (1996). Induction of embryogenesis with colchicine instead of heat in microspores of *Brassica napus* L. cv. Topas. *Planta* 198, 433–439. doi: 10.1007/BF00620060

Zhou, W. J., Hagberg, P., and Tang, G. X. (2002). Increasing embryogenesis and doubleing efficiency by immediate colchicine treatment of isolated microspores in spring Brassica napus. *Euphytica* 128, 27–34. doi: 10.1023/A:1020687517888

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Yuan, Su, Liu, Li, Fang, Yang, Zhuang, Zhang, Lv and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 June 2016
doi: 10.3389/fpls.2016.00943

An Efficient Method for Adventitious Root Induction from Stem Segments of *Brassica* Species

*Sandhya Srikanth, Tsui Wei Choong, An Yan, Jie He and Zhong Chen**

Natural Sciences and Science Education, National Institute of Education, Nanyang Technological University, Singapore, Singapore

Plant propagation via *in vitro* culture is a very laborious and time-consuming process. The growth cycle of some of the crop species is slow even in the field and the consistent commercial production is hard to maintain. Enhanced methods of reduced cost, materials and labor significantly impact the research and commercial production of field crops. In our studies, stem-segment explants of *Brassica* species were found to generate adventitious roots (AR) in aeroponic systems in less than a week. As such, the efficiency of rooting from stem explants of six cultivar varieties of *Brassica* spp was tested without using any plant hormones. New roots and shoots were developed from *Brassica alboglabra* (Kai Lan), *B. oleracea* var. acephala (purple kale), *B. rapa* L. ssp. *chinensis* L (Pai Tsai, Nai Bai C, and Nai Bai T) explants after 3 to 5 days of growing under 20 ± 2°C cool root zone temperature (C-RZT) and 4 to 7 days in 30 ± 2°C ambient root zone temperature (A-RZT). At the base of cut end, anticlinal and periclinal divisions of the cambial cells resulted in secondary xylem toward pith and secondary phloem toward cortex. The continuing mitotic activity of phloem parenchyma cells led to a ring of conspicuous white callus. Root initials formed from the callus which in turn developed into ARs. However, *B. rapa* var. *nipposinica* (Mizuna) explants were only able to root in C-RZT. All rooted explants were able to develop into whole plants, with higher biomass obtained from plants that grown in C-RZT. Moreover, explants from both RZTs produced higher biomass than plants grown from seeds (control plants). Rooting efficiency was affected by RZTs and explant cuttings of donor plants. Photosynthetic CO_2 assimilation rate (A_{sat}) and stomatal conductance (g_{ssat}) were significantly differentiated between plants derived from seeds and explants at both RZTs. All plants in A-RZT had highest transpiration rates.

Keywords: aeroponics, *Brassica*, explants, root zone temperature, rooting, stem-segment

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Abu Hena Mostafa Kamal,
University of Texas at Arlington, USA
Appakan Shajahan,
Jamal Mohamed College, India

****Correspondence:***
Zhong Chen
zhong.chen@nie.edu.sg

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *22 December 2015*
Accepted: *13 June 2016*
Published: *29 June 2016*

Citation:
Srikanth S, Choong TW, Yan A, He J and Chen Z (2016) An Efficient Method for Adventitious Root Induction from Stem Segments of Brassica Species. Front. Plant Sci. 7:943. doi: 10.3389/fpls.2016.00943

INTRODUCTION

Green leafy vegetables, known for their high nutritional content, are consumed by humans for good health and dietary benefits. The Brassicaceae family encompasses the Brassiceae tribe, which includes wide varieties of agriculturally and economically significant species. The genus *Brassica*, including kales, cabbages, broccoli, cauliflower, brussel sprouts, and kohlrabi, comprise of biennially herbaceous plants classified by their characteristic morphology of edible parts

(Parkin et al., 2014). *B. alboglabra,* also known as Chinese kale or Kai Lan, is among the ten most marketable vegetables in Southeast Asian countries, such as Hong Kong, Thailand and China (Rakow, 2004). *B. oleracea* var. *acephala* (purple curly kale), related to the common cabbage, is a biennial temperate crop that is cultivated as an annual. Their flower buds and leaves are used as potherbs or greens (Velasco et al., 2007). Mizuna, a cultivated variety of *B. rapa* var. *nipposinica*, is a leafy vegetable commonly found in Japanese salads. Pai Tsai and Nai Bai are the cultivated varieties of *B. rapa* L. ssp. *chinensis* that is gaining popularity in Western menus where they are often steamed or stir fried (Rochfort et al., 2006). *Brassica* species play a significant role in agriculture and horticulture fields and contribute significantly to economies and population health worldwide (Zhao, 2007). Additionally, these *Brassica* species also represent an excellent system for studying numerous aspects of plant biology (Cheng et al., 2011).

Seed companies and industries face many challenges to acquire seeds from *Brassica* cultivars as it takes an especially long time to obtain. Conventional methods of hybrid seed production involve selling inbred lines for at least ten generations, while the development of homozygous plants of anther culture takes at least a year (Yang et al., 1992). Further, unlike other crops, Brassica seed production is limited to the fields with a minimum 5-year gap between seed crops and at least 2 years exclusion of any Brassica species in the same field (Ellis, 2007). Hence, a quick alternative in the commercial field would be to use vegetative propagation via stem cuttings. This minimizes seed usage while allowing for new plants to be developed within a week shorter time interval.

Vegetative propagations, such as leaves, cladode, stem or branch cuttings, are well-known methods of asexual propagations. In general, stem cutting is the most popular method of propagation for commercial plantings worldwide. However, operating costs are high as a continuous supply of fresh materials, such as peat moss, vermiculite, coir pith, root trainers, and fungicides, are required for the existing stem cutting propagation method. Rooting hormones, such as auxin, are also required for *de novo* root formation in explants. Though auxin stimulates root initiation, it also habitually leads to callus formation and expression of genes that are not necessarily related to root initiation (Welander et al., 2014). Since ARs may originate independently and directly from explant tissue rather than from callus (Fink, 1999), a more efficient rooting method with little or no callus formation would be desirable. Thus, a new method that could increase the speed of propagation whilst lower propagation costs would be an ideal alternative approach.

Soilless culture systems are useful for both research and commercial applications for food crops. An example is an aeroponic system which allows for plants to grow whilst their roots are suspended in air. As previous studies (Weathers and Giles, 1988; Zobel, 1989; Luo et al., 2012) have suggested that aeroponics is the optimum system for growing intact plants or excised roots and tissue cultures, this research explores the possibility of vegetative propagation of temperate *Brassica* species (Asian greens) in an aeroponic system within a tropical greenhouse, with the manipulation of only RZTs He et al. (2013) and He (2014) and in the absence of any hormonal applications. Therefore, this study describes a rapid and efficient rooting and whole plant regeneration methodology for six *Brassica* species. This method of root and shoot development using aeroponics can be applicable to all commercial *Brassica* cultivars. The findings of this study could be applied in the mass propagation of vegetable crops, shortening the growth cycle as seed germination and seedling development periods can be eliminated.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The six cultivars of *Brassica* spp used in the study were: Kai Lan (*B. alboglabra*), purple curly kale (*B. oleracea* var. *acephala*), Mizuna (*B. rapa* var. *nipposinica*), Pai Tsai, Nai Bai C, and Nai Bai T (*B. rapa* L. ssp. *chinensis*). The seeds were germinated at 25°C room temperature on wet tissue paper in petri plates for 3 days. Seedlings were inserted into polyurethane cubes and acclimatized at 35°C in the greenhouse for 3 days before transplanting into the aeroponic troughs at 20 ± 2°C cool root zone temperature (C-RZT) and 30 ± 2°C ambient root zone temperature (A-RZT), respectively. The roots of all plants were misted with full strength (pH 6.8, EC 2.2 mS) Netherlands Standard Composition (Douglas, 1982) nutrient solution for 1 min at 5 min intervals. All these conditions of aeroponic troughs were maintained constant throughout the experiment. Midday shoot temperatures varied between 35 to 40°C. After 45 days of transplanting, at the time of harvesting stage, few of the stem segment explants with one leaf were transplanted in their respective A- and C-RZTs to raise second generation vegetables. Another set of the seeds was germinated provided with a same nutrient solution, on the same day of explants transplantation as controls to compare the growth performance between explants and controls. Explants started rooting and established few short roots at the time of control seedling establishment in both RZTs. Hence, the establishment time period (1 week) is a prerequisite for both explants and control seeds for rooting and seed germination/seedling development, respectively. However, explants showed vigorous growth as compared to control seedlings in later stages of development.

Plant Histology

Explants of all six varieties of *Brassica* species followed same rooting orientation in a circular mode toward the parenchymatous cortex region (**Figure 1**). Therefore, as an example tender and small stems of Kai Lan explant samples were collected during the callus formation and root initiation and fixed them in 10% formaldehyde for 3 days before being dehydrated through graded solutions of alcohol, cleared in xylene, and infiltrated with wax. Embedded sections were transversely and longitudinally sectioned using Leica rotary microtome at 7 μm thicknesses and transferred onto glass slides (after spreading well in a water bath at 45°C). The slides were dried on the hot plate at 45°C for 5 min. Completely dried and water free slides were dipped twice in xylene for 5 min each, twice in 100% ethanol for 5 min, once in 95% ethanol for 1 min and again



FIGURE 1 | The comparable orientation of adventitious rooting in a circular form toward the parenchymatous cortex in the cut ends of six *Brassica* sp explants. (A) Kai Lan. **(B)** Purple curly kale. **(C)** Mizuna. **(D)** Pai Tsai. **(E)** Nai Bai C. **(F)** Nai Bai T.

once in 70% ethanol for 1min. The slides were then air dried and stained in 0.05% (w/v) toluidine blue in distilled water for 2 min, rinsed three times with deionized water to remove excess stain. The slides were then dried and mounted using Permount^{TM} mounting medium (ProSciTech, Thuringowa, Australia). Root generation from stem segments of Kai Lan was examined by using Olympus BX60 microscope.

Measurements of Leaf Photosynthetic Pigments

Samples were collected from explants and controls after 6 weeks of their establishment in both RZTs. 0.05 g of fresh leaves of three plants of each variety were soaked in 5 ml N, *N*-dimethylformamide and left in the dark at 4°C for 48 h. Absorption at wavelengths of 480 nm, 647 nm, and 664 nm were measured using a spectrometer (UV-2550, Shimadzu, Japan) and concentrations of chlorophyll a, chlorophyll b and carotenoids were then calculated (Wellburn, 1984).

Measurements of Photosynthetic Parameters

Newly expanded leaves of intact plants were analyzed for photosynthetic parameters such as light saturated photosynthetic CO_2 assimilation rate (A_{sat}), stomatal conductance (g_{ssat}), intercellular CO_2 concentration (C_i) and transpiration rate using an open infrared gas analysis system (LI-COR), 6 weeks after the establishment of both explants and controls in C-RZT and A-RZT. The LED light source was set at a photosynthetic photon flux density of 1000 μmol m^{-2} s^{-1}. Leaf chamber temperature, relative humidity, and average ambient CO_2 concentration were 29°C, 70% and 396 $\pm$ 3 μmol mol^{-1}, respectively.

Measurements of Growth, Productivity, and Water Content

Matured explants and controls were harvested 45 days after transplant and germination, respectively, to obtain shoot and root fresh weights (FWs) and dry weights (DWs). For FW, the roots of plants were blotted with a tissue to remove excess water prior to weighing. DW was determined after drying the same plant samples at 100°C for 72 h. The means for each cultivar was determined from three plants. Morphological traits, such as leaf area, and shoot and root lengths, were also photographed and measured. Leaf area was measured using WinDIAS3 v3.2.1 (Delta-T Devices Ltd). Water content (WC) was calculated as follows: WC = [(FW−DW)/FW] × 100%.

RESULTS

The Efficiency of Different *Brassica* Vegetable Explants to Generate Roots

Stem cuttings/explants of 20 for each mode of cutting from lower, middle, and top portions (**Figure 2A**) of the donor plants were tested for the rooting ability in both RZTs to avoid highest mortality rate in explants propagation (**Table 1**). During this investigation, the highest rooting ability was found in all explants with shoot apical meristem (SAM) that grew under C-RZTwhich can be attributed to the coordination between SAM and root development in C-RZT. All explants rooted in the same level under A-RZT except Mizuna explants which died due to exposure to the highest midday light intensity up to 1000 μmolm^{-2}s^{-1} and air temperature $\leq$40°C. KaiLan, Nai Bai C, and Nai Bai T explants from the middle and lower portions of their respective donor plants were able to root normally and developed new shoots from axillary meristems. Whereas, Pai Tsai and kale explants from middle and lower portions were unable to generate new roots and shoots and eventually died after 2–3 days of planting on both C-RZTand A-RZT.

De novo Adventitious Root Formation in Aeroponics

Soon after planting the stem segment explants on 20 $\pm$ 2°C and 30 $\pm$ 2°CRZTs in tropical Aeroponics, the outer layer of wounded cells died and later formed a necrotic dark layer. This layer helped to protect the cut surface from desiccation and

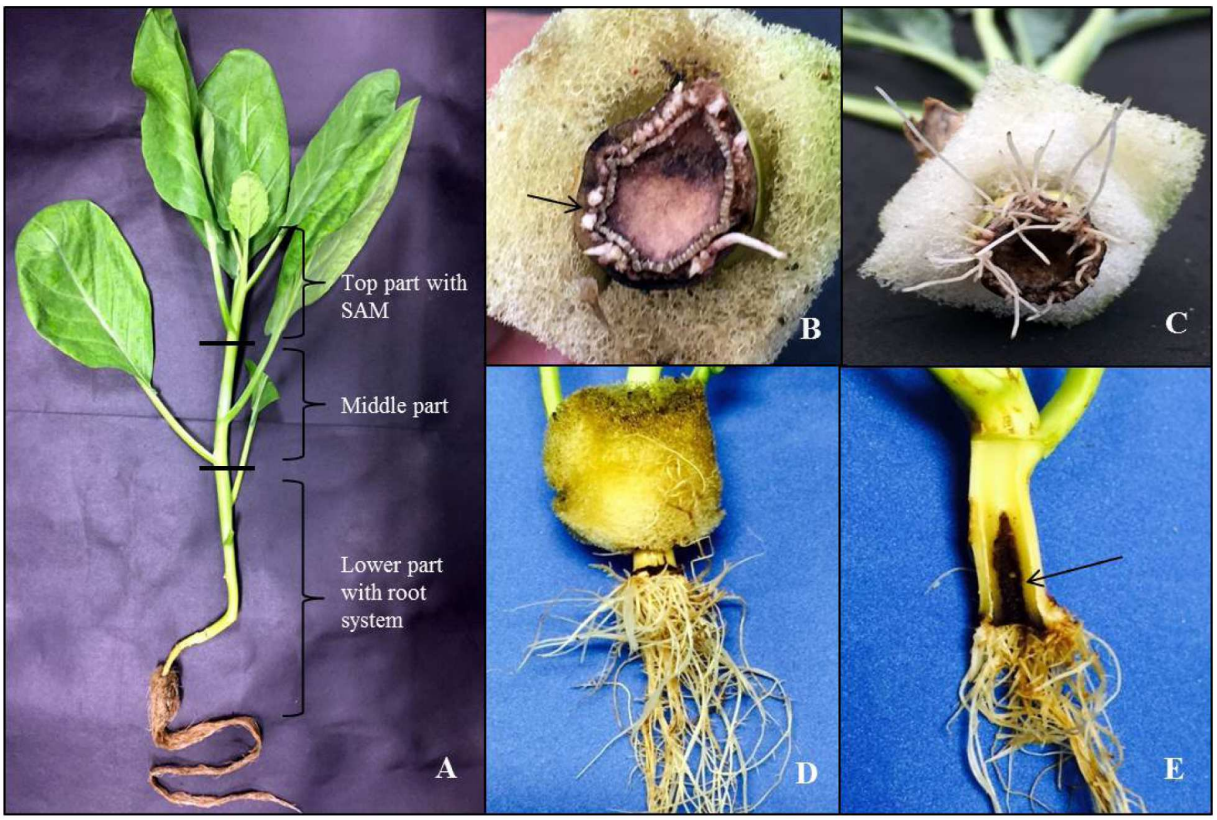


FIGURE 2 | Early events of callus formation and elongation of adventitious roots (ARs) in Kai Lan stem explants. (A) Kai Lan plants with cutting regions and explants positions. **(B)** Arrow pointing to the callus initiation at the cut end. **(C)** Roots formed in a circular ring. **(D)** Root elongation. **(E)** Arrow showing the longitudinal section of fresh regenerating stem showing degenerating central pith region.

pathogens attack. Living cells underneath this dark layer has begun divisions after 2 days of wounding and an outer layer of parenchyma cells after rapid mitotic divisions formed a circular mass of cells which later formed little white undifferentiated tissue called callus (**Figure 2B**). At the base of cut end, anticlinal and periclinal divisions of the cambial cells resulted in secondary xylem toward pith and secondary phloem toward cortex (**Figure 3**). The xylem vessels became a thick walled to ease the rapid hydraulic conductivity for regenerating explants (**Figures 3C,D**).

TABLE 1 | Stem segment ex-plant rooting ability in six *Brassica* species *B. alboglabra* (KaiLan), *B. oleracea* var. *acephala* (purple kale), *B. rapa* L. ssp. *Chinensis* L (Pai Tsai, Nai Bai C, and Nai Bai T).

Cultivar name	Lower stem with root system	Middle stem cuttings	Shoot apical meristem
Kai Lan	✓	✓	✓
Mizuna	✓	–	✓
Kale	✗	✗	✓
Pai Tsai	✗	✗	✓
Nai Bai C	✓	✓	✓
Nai Bai T	✓	✓	✓

The newly formed phloem parenchymatous cells after vigorous cell divisions formed root callus growth proliferations (**Figures 4, 5,** and **6A**). Root initials formed from the callus in the vicinity of the vascular cambium and phloem ray parenchyma (**Figure 6A**), which had become meristematic by dedifferentiation. Further cell divisions occurred and the meristematic area had become more organized with the formation of a root initial (**Figures 2** and **6**). Ultimately, these root initials developed into an organized root primordia in the secondary phloem and cortex (**Figure 6A**). Later, the root primordia grew outwardly through stem tissues and formed the vascular tissue connections between the primordia and vascular tissue of stem cutting. Upon emergence from the stem segment, the ARs have already developed a root cap as well as a complete vascular connection with the originating stem (**Figures 6D–H**). Eventually, parenchymatous cells of the cortex also contributed to root formation covering the entire base with many roots and root initials (**Figures 5B** and **6H**). Visible root initials emerged from the cuttings after 3–5 days of planting in C-RZT aeroponics whereas roots were visible only after 4–7 days of planting inA-RZT. Once primordia are formed, there was a comparable time period of 5–7 days between root primordia elongation (emergence) and maximum rooting in both RZTs. Even in the C-RZT, some of the

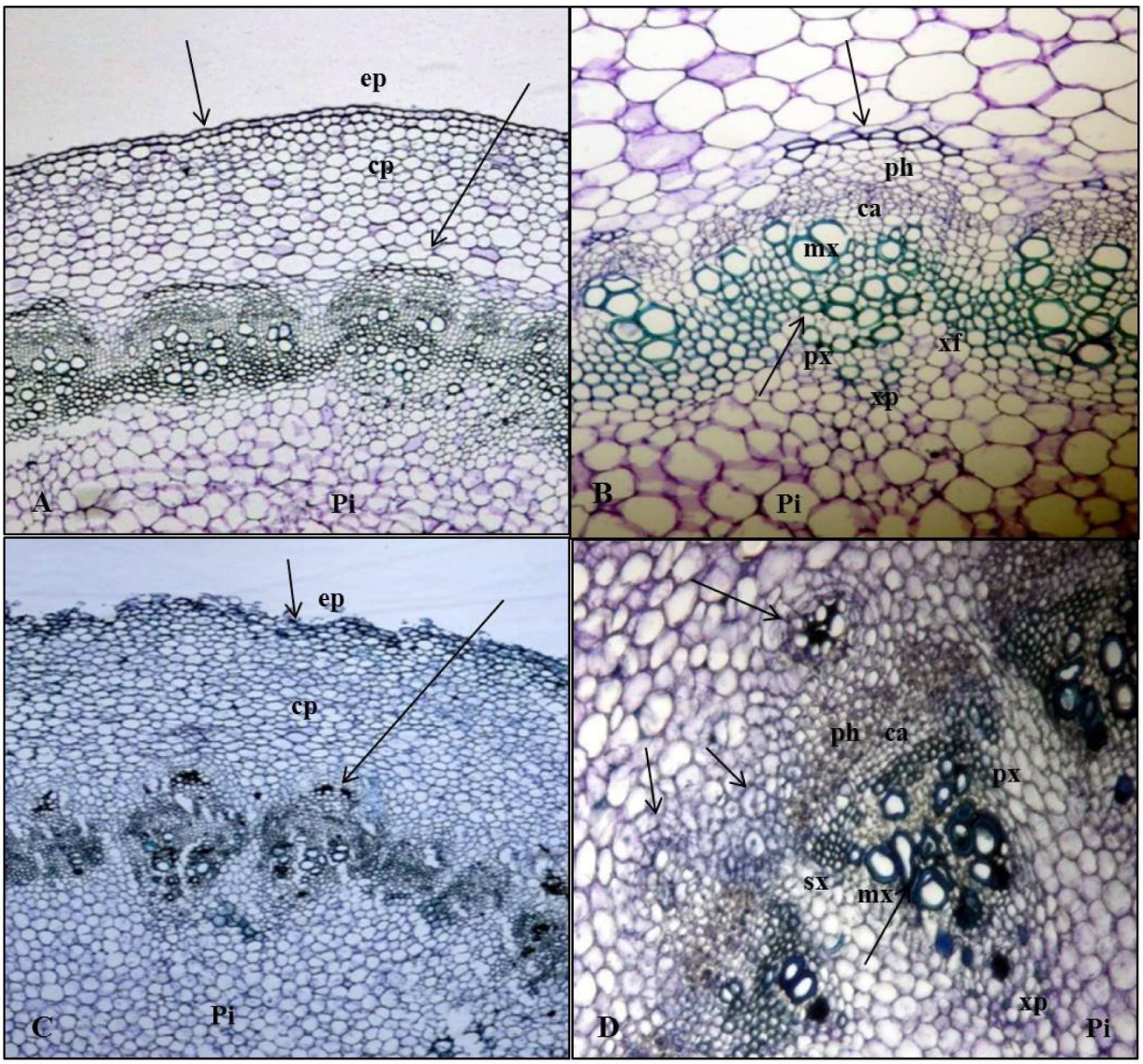


FIGURE 3 | Anatomy of AR formation from phloem parenchyma in Kai Lan stem explants of control (A,B-arrows showing normal tissues) and rooted (C,D-arrows showing the abnormal tissues and cell division) samples. (A) Control stem transverse section showing epidermis (ep), cortex parenchyma (cp) and pith (pi) in normal tissue orientation. **(B)** Control stem showing vascular bundle (phloem (ph), cambium (ca), metaxylem (mx), protoxylem (px), xylem parenchyma (xp), and pith (pi) with aregular arrangement of cells. **(C)** Rooted stem transverse section showing irregular epidermis (ep) and cortex parenchyma (cp) with an abnormal tissue orientation. **(D)** The continuing mitotic activity of phloem parenchyma cells led to a ring of conspicuous meristematic tissue complexes evident in the cortex region of the rooted stem section. Vascular bundle showing the secondary xylem (sx) and phloem (ph) cells derived from active cambial cells (ca).

purple kale cuttings were delayed rooting even after 7 days of planting. This delay was due to variability in cuttings from different-sized stock plants but once root primordia formed, root emergence consistently occurred within a week period in both RZTs.

Explant Establishment, Growth, and Productivity

Among six vegetable explants planted on C- and A-RZTs, Kai Lan, Nai Bai C, and Nai Bai T showed the highest survival rate as compared to other three vegetable explants such as Kale, Pai Tsai, and Mizuna. In C-RZT, Kai Lan explants exhibited 95% survival rate among all the Kai Lan cuttings planted, whereas explants in A-RZT showed 80% survival rate and 20% mortality rate. However, Nai Bai C and Nai Bai T explants showed the highest survival rate as 96 and 98% in C-RZT and 90 and 93% survival rate in A-RZT, respectively. Both Mizuna and Kale explants showed 60% survival rate in C-RZT and only Kale explants survived up to 45% in A-RZT. While the survival rate of Pai Tsai explants was 75% in C-RZT and 70% in A-RZT. In C-RZTs, control seedlings from six vegetable varieties established

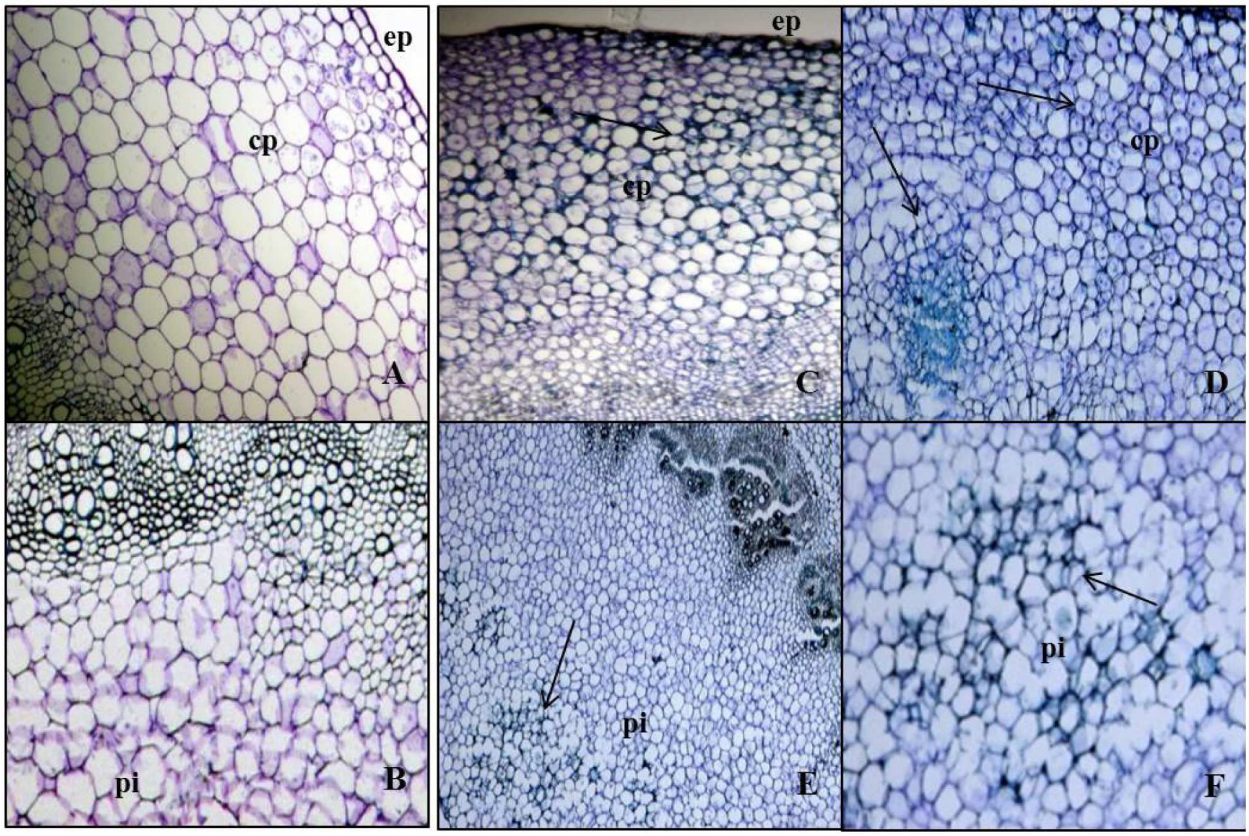


FIGURE 4 | Cell divisions in cortex and deterioration in pith occurred post root formation from the secondary phloem. (A) Transverse section of Kai Lan control stem showing normal epidermis (ep) and cortex parenchyma (cp) regions. **(B)** Control stem showing normal pith at the center. **(C)** Transverse section of rooted stem showing active mitotic divisions in the cortex to initiate meristematic region. **(D)** Rapid mitotic divisions resulted in a circular mass of tissue in cortex parenchyma (cp). **(E)** Rooted stem showing early events of deterioration in central pith (pi). **(F)** Enlarged picture showing the cell disintegration in central pith (pi) (arrows pointing toward the cell division).

successfully, exhibiting up to 98% survival rate in Kai Lan, Nai Bai C, and Nai Bai T, 95% in Mizuna and Kale, whereas 96% in Pai Tsai. But in A-RZT, all six varieties showed 1 to 2% less establishment rate as compared to C-RTZ. Since all explants/cuttings used in this experiment were almost uniform in each species, the number of growth measurements such as shoot length, root length, leaf area, fresh and DWs of shoots and roots were recorded and analyzed.

Application of different RZTs significantly affected the plant growth components, especially shoot and root lengths. Explants grown in both RZTs showed the highest biomass as compared to their controls. Nevertheless, explants of C-RZT were large stature with increased biomass compared to the explants of A-RZT. While this difference in the plant stature and biomass was also observed in the control plants of both RZTs. Moreover, all results indicating clearly that the controls were significantly smaller with very low biomass when compared with the explants in both RZTs (**Table 2**; **Figures 7** and **8**). Kai Lan explants showed almost 12.4 cm shoot length and 23.5 cm root length difference from C- to A-RZTs (**Table 2**). This larger difference in Kai Lan plant stature was also correlated with twofold decreased leaf area in A-RZT (**Figure 7**). Although there was not much difference in shoot length of Mizuna explants and controls, it was evident from the results that Mizuna explants grew well with more leaves in the C-RZT and showed almost threefold increased leaf area. Whereas kale, Pai Tsai, Nai Bai C, and Nai Bai T explants showed low to moderate differences in the plant heights and leaf area measurements, but still explants of C-RZT were significantly ($^{*}p < 0.05$, $^{**}p < 0.01$) superior to the A-RZT (**Table 2**, **Figure 7**).

Total fresh and DWs of shoots and roots were determined from a single harvest of three plants from each variety when explants reached harvest maturity. In correlation with the plant stature, the biomass (FW and DW) of all vegetable explants grown under cool and A- RZTs showed the highest and remarkable increase ($^{*}p < 0.05$, $^{**}p < 0.01$) compared to the control plants (**Figure 8**). However, FW and DW were significantly affected by imposing different RZTs. Among all six vegetable varieties, Kai Lan explants from C-RZT showed almost twofold increased FW in concurrence with their large plant stature. However, from the data, it was evident that the DW of all vegetables decreased drastically as they possessed a high

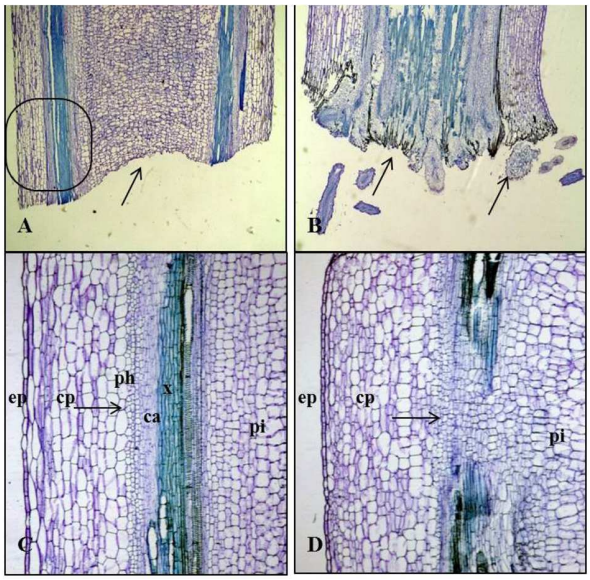


FIGURE 5 | Longitudinal sections of Kai Lan control and rooted stems showing variation in tissue orientation. (A) Longitudinal section of the control stems (arrow indicating the blunt end without any divisions). **(B)** Longitudinal section of rooted stem showing rooting from basal wound tissue (arrow indicating the cell divisions at the cut end). **(C)** Enlarged picture of the control stem longitudinal section showing epidermis (ep), cortex parenchyma (cp), the arrow showing the regular phloem (ph), cambium (ca) xylem (x) and pith (pi). **(D)** Enlarged picture of rooted stem L.S showing epidermis (ep), cortex parenchyma (cp), pith (pi) and an arrow showing the irregular vascular bundles with rapidly dividing cells.

amount of water content in the root and shoots. Explants of Kai Lan shoots and roots showed 93% WC in the C-RZT whereas 92% in A-RZT. Mizuna explants possessed almost 91% WC in both roots and shoots, whereas controls showed 92% WC in both RZTs. Kale explants showed almost 90% WC in shoots and roots in both RZTs. While shoots of kale controls contained with 90.5% in C-RZT and 87% in A-RZT. But kale control roots retained 92.8% WC in A-RZT and only 84% in C-RZT. Whereas, all Pai Tsai, NaiBai C, and NaiBai T plants (both explants and controls) showed 95% WC in shoots and up to 91% WC in roots of both RZTs. From the results, it was clear that almost all plants retained more than 90% WC.

Photosynthetic Pigments and Performance

Kai Lan explants had lower carotenoid content (**Figure 9C**) in C-RZT compared to A-RZT. In C-RZT, Mizuna explants had higher chlorophyll a/b ratio (**Figure 9A**), total chlorophyll (**Figure 9B**) and carotenoid content than its control plants. For Nai Bai C, there was higher total chlorophyll and carotenoid content in C-RZT for the control plants, but the reverse was observed for its explants. Such similar results were also observed in Nai Bai T, with its chlorophyll/ carotenoid ratio (**Figure 9D**) being lower for both control and explant. However, the chlorophyll/carotenoid was similar for both control and explants of Nai Bai C. A_{sat} (**Figure 10A**) was higher for all Kai Lan plants in C-RZT, but similar between control and explants in both RZTs for the rest of the plant types. g_{ssat} (**Figure 10B**) was generally higher for most plants in C-RZT with the exception of

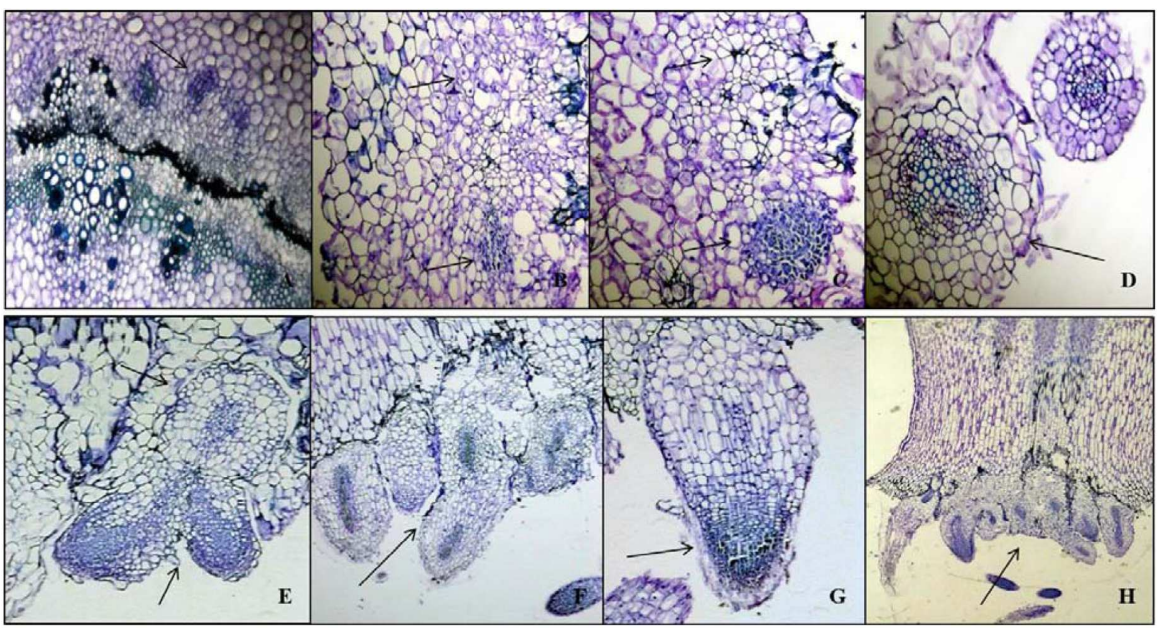


FIGURE 6 | Histological characteristics of AR development in Kai Lan stem explants. (A–D) Transverse sections showing early events of mitotic divisions and callus initiation. **(A)** Vascular bundle showing thick-walled vascular tissues (Arrow pointing to the callus initiation from phloem parenchyma). **(B)** Arrows showing the mitotic cell divisions. **(C)** Arrows indicating the callus initials. **(D)** Arrow showing the round callus. **(E–H)** Arrows in the longitudinal sections showing root primordial initiation from callus **(E,F)**, root cap development **(G)** and AR formation from meristematic region of explant cut ends **(H)**.

TABLE 2 | Explants and controls showed a significant difference (*$p < 0.05$, **$p < 0.01$; $n = 10$) in their shoot and root length under C- and A-RZTs.

Plant sample	C-RZT		A-RZT	
	Shoot length (cm)	Root length (cm)	Shoot length (cm)	Root length (cm)
Kai Lan explant	45.5 ± 1.3**	74.8 ± 3.6*	33.1 ± 2.4*	51.3 ± 1.2**
Kai Lan control	24.6 ± 0.6	51.3 ± 4.3	22.7 ± 2.6	43.6 ± 1.0
Mizuna explant	34.0 ± 2.6	83.3 ± 4.7**	–	–
Mizuna control	30.0 ± 0.5	59.5 ± 1.3	23.9 ± 1.6	28.3 ± 3.8
Kale explant	37.4 ± 2.8**	69.0 ± 4.7**	25.9 ± 4.2*	41.0 ± 1.0
Kale control	16.0 ± 1.7	38.3 ± 0.6	12.0 ± 0.5	36.8 ± 3.5
Pai Tsai explant	27.0 ± 1.1	80.0 ± 6.2*	23.3 ± 2.6	45 ± 1.5
Pai Tsai control	22.0 ± 1.6	51.3 ± 4.4	22.6 ± 1.5	38.3 ± 4.4
Nai Bai C explant	15.5 ± 0.8	61.3 ± 3.6	16.3 ± 2.7	59.6 ± 4.3*
Nai Bai C control	11.1 ± 1.6	55.0 ± 5.5	10.3 ± 0.3	29.6 ± 3.7
Nai Bai T explant	18.5 ± 1.8	69.1 ± 6.7**	14.5 ± 1.0*	51 ± 4.5
Nai Bai T control	15.1 ± 0.6	42.0 ± 3.7	10.5 ± 0.5	36.6 ± 3.6

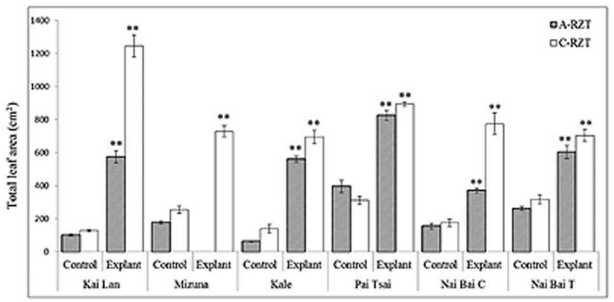


FIGURE 7 | Total leaf area of explants and controls of six cultivar varieties of *Brassica* spp. grown under C-RZT (20 ± 2°C) and A-RZT t (30 ± 2°C) in a tropical aeroponics greenhouse for 45 days. Each bar graph is the mean of 3 measurements from three different plants ($n \geq 3$). Vertical bars represent standard errors. Significance: **$p < 0.01$.

Kai Lan which had lower g_{ssat}. Transpiration rate (**Figure 10D**) was generally lower in C-RZT, than A-RZT, for all plants except for Nai Bai C which had similar rates at both RZTs. Nai Bai T had higher g_{ssat} (**Figure 10B**) and C_i (**Figure 10C**), though a significantly lower transpiration rate for both control and explants, at C-RZT.

DISCUSSION

Plant tissues have the enormous regeneration capacity, and entire plant can be developed from a single cell or small cuttings/explants (Xu and Huang, 2014). The AR formation is of great importance for vegetative propagation, but difficult to achieve in many crop species. In the present study, ARs were successfully developed from stem-segment explants of *Brassica* species on tropical aeroponics. *De novo* induction of roots from stem cuttings in plants involve the induction of meristems from adult somatic cells that are not determined to originate a meristem in normal development. Usually, AR formation is induced in stem cuttings, which experience a stimulus, such as wounding (Abarca and Díaz-Sala, 2009). Complex cellular processes involved in the AR formation are cell reorganization, induction of cell divisions, the organization of a root primordium and root development and emergence (Legué et al., 2014).

The AR formation was direct; i.e., in an organized mode without a long intervening period of callus formation. The roots grew through the cortex and often emerged out from a small round callus during the 3 to 7 day's period after inserting the cuttings into aeroponic boards (in both RZTs). Interestingly, before root emergence, we have observed a ring of secondary vascular tissues developed from the actively dividing cambial cells at the wound site. During the rooting, this kind of tissue differentiation was observed only at the cut end and not above the wound region. Hence, it is emphasized that the wound signaling and continuous nutrient solution mist had led to the immediate auxin accumulation at the cut end which induced few small calluses and then rooting. At the base of cut end, anticlinal and periclinal divisions of the cambial cells resulted in secondary xylem toward pith and secondary phloem toward cortex. The continuing mitotic activity of secondary phloem parenchyma cells led to a ring of conspicuous white meristematic tissue complexes called 'callus'. Root initials formed from the callus which in turn developed into ARs in the vicinity of the vascular cambium and phloem ray parenchyma. The study highlights that the hormone free cuttings can produce roots at multiple positions around the vascular tissue and so this propagation method can produce more ARs at the base of each cutting which resulted in higher survival and growth rate of explants.

In this experiment, significant results have been recorded for stem cutting propagation which would be demonstrating the possibility and success rate of vegetative propagation in tested *Brassica* samples and other vegetable species. The results are also representing the significant difference in the biomass of explants and controls. Among all, Kai Lan explants showed more remarkable FW and DW in cool RZT compared with other explants and controls. Since, explants were collected from the matured harvest stage donor plants, their well-established

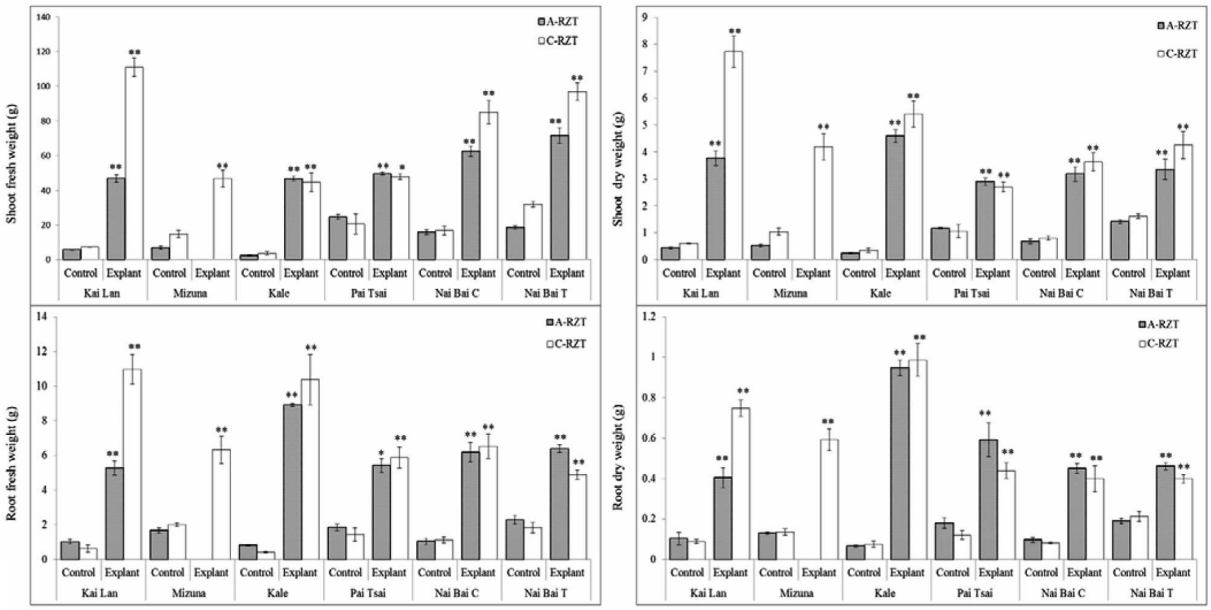


FIGURE 8 | Fresh and dry weights (DWs) of shoots and roots of explant and control plants. Six cultivar varieties of *Brassica* spp. grown under C-RZT (20 ± 2°C) and A-RZT (30 ± 2°C) in a tropical aeroponics greenhouse for 45 days. Each bar graph is the mean of three measurements from three different plants ($n \geq 3$). Vertical bars represent standard errors. Significance: $^{*}p < 0.05$, $^{**}p < 0.01$.

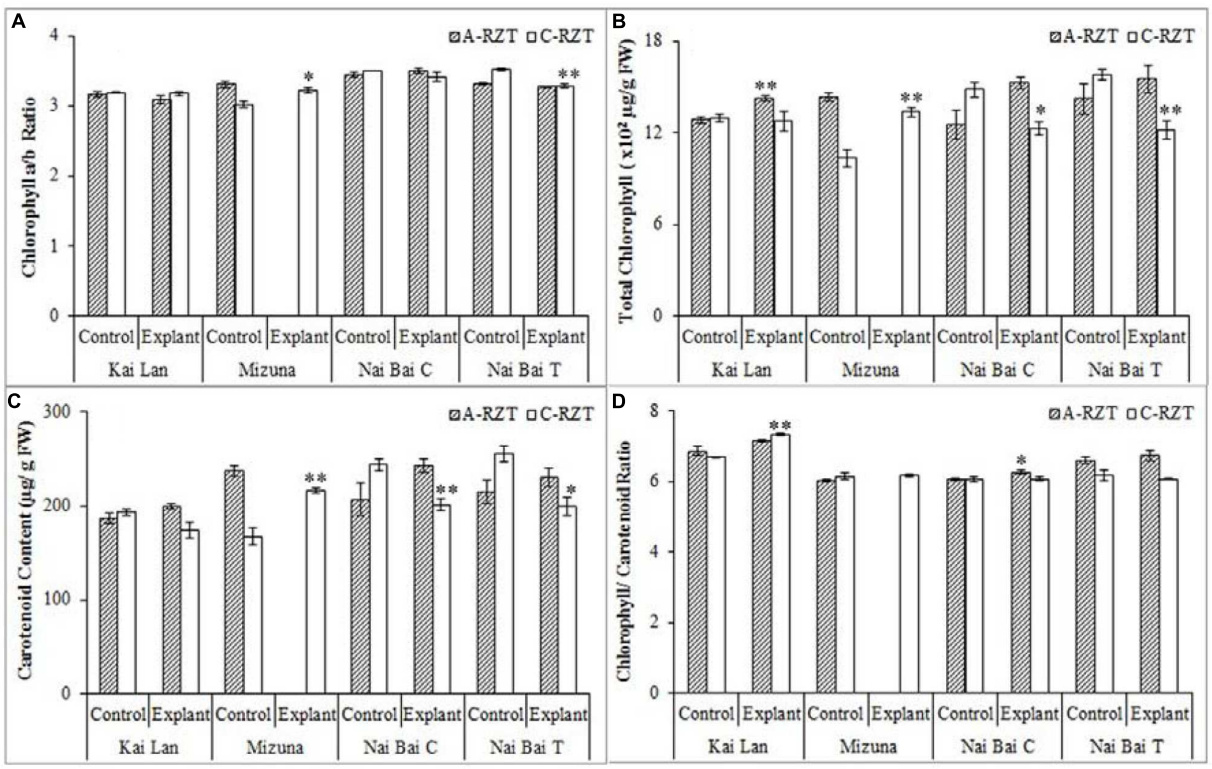


FIGURE 9 | Chlorophyll a/b ratio (A), total chlorophyll concentration (B), carotenoid concentration (C), and chlorophyll/ carotenoid ratio (D) of Kai Lan, Mizuna, Nai Bai C, and Nai Bai T grown in A-RZT and C-RZT. Each bar graph is the mean of three measurements from three different plants ($n \geq 3$). Vertical bars represent standard errors. Significance: $^{*}p < 0.05$, $^{**}p < 0.01$.

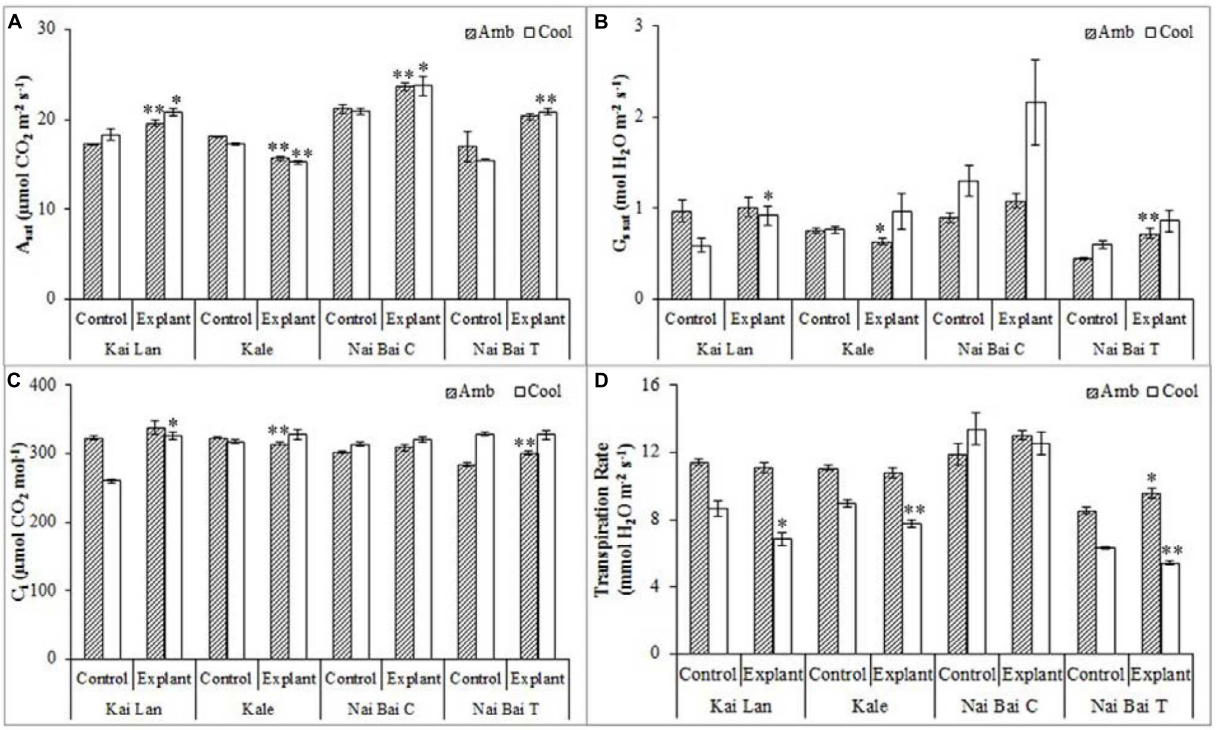


FIGURE 10 | Photosynthetic parameters such as light saturated photosynthetic CO_2 assimilation rate (A_{sat}) (A), stomatal conductance (g_{ssat}) (B), intercellular CO_2 concentration (C_i) (C), and transpiration rate (D) of Kai Lan, Mizuna, Nai Bai C, and Nai Bai T grown in A-RZT and C-RZT. Each bar graph is the mean of three measurements from three different plants ($n \geq 3$). Vertical bars represent standard errors. Significance: $^*p < 0.05$, $^{**}p < 0.01$. Amb: ambient temperature; Cool: cool temperature.

meristematic regions, and continuous nutrient supply through root zone area of aeroponics could be contributed for their quick recovery and root formation in a few days' period. Once explants established the root system, they started growing vigorously and resulted in large plant stature. Meanwhile, seeds germinated on the same day of the explants planting took 2–3 days for germination and then 3 days for the establishment. But even after transplanting on the aeroponic systems, seedlings grew very slowly in the first few days until they established potential meristematic regions and root systems to support rapid growth. Eventually, seedlings of all vegetable varieties (Kai Lan, Mizuna, Kale, Pai Tsai, Nai Bai C, and Nai Bai T) started growing rapidly after one month of transplanting. While control plants reached halfway to the harvest stage, explants were grown up to harvest stage. This resulted in the high biomass content (FW and DW) of the roots and shoots of explants compared to control plants. While total leaf area was influenced by C-RZTs conditions, as there were more leaf number and specific leaf area was recorded little higher than A-RZT. However, explants recorded highest SLA compared to control plants due to changes in leaf thickness and leaf density. This simple and reliable method of explants regenerate into a whole plant, thereby rapid vegetative propagation is suitable for the vegetable varieties tested in both RZTs except Mizuna in A-RZT. But still, Mizuna showed significantly increased biomass when compared to its control plants in C-RZT. Even though Kale explants showed higher biomass content, they encountered a poor establishment problem in A-RZT. While other vegetables such as Pai Tsai, Nai Bai C, and Nai Bai T had shown greater establishment rate and yielded more in both RZTs. C-RZT also promoted the higher shoot and root lengths of Kai Lan, Mizuna, Kale, Pai Tsai, Nai Bai C, and Nai Bai T explants, suggested that plants grown in C-RZT possess a higher water and nutrition uptake capacity, which would contribute to the productivity of plants. The ratio of leaf chlorophyll a and chlorophyll b contents were not significantly different among controls and explants in both RZTs. However, total chlorophyll content is slightly higher in most of the plants in A-RZT. Nevertheless, total chlorophyll/ carotenoid ratio found to be almost similar in controls and explants of both RZT, except Nai Bai which showed a little higher ratio in A-RZT. These results demonstrating that the photosynthetic pigments contents are species specific and their performance is unique in each species. Photosynthetic parameters such as light saturated photosynthetic CO_2 assimilation rate (A_{sat}) is considerably higher in almost all explants at both RZTs, while stomatal conductance (g_{ssat}) is higher in C-RZT plants. Inter cellular CO_2 concentration (C_i) is not much influenced by different RZTs and is relatively uniform among all plants. Whereas, transpiration rate is significantly higher in A-RZT plants compared to C-RZT plants except Nai Bai. Despite all variations in photosynthetic parameters, all explants were healthy and well advanced in growth in comparison to control plants. The above discussion is

also demonstrating that the effect of C-RZT and A-RZT is again species-dependent.

In the present study, by using the aeroponic systems we just studied the AR formation and successful establishment of the stem segment explants of *Brassica* vegetable species without using any plant hormones, which would be useful in the vegetative propagation of leafy vegetable crops that has huge demand worldwide. Moreover, there is a need to develop different farming systems to secure the continuous vegetable production in space limited cities such as Singapore. By adapting new cultivating techniques like vegetative propagation on aeroponic farming systems, constant leafy vegetable supply is conceivable and compels the vegetable import in urban areas. There may be other instances in which aeroponic vegetative propagation can be used as an alternative to seed propagation. They include easy and rapid multiplication of selected genotypes which are generated from the conventional breeding program or induced variants from cells, tissue, or organ culture, genetic transformation, propagation of parents for hybrid seed production and speedy propagation of asexually propagated crops. Aeroponic propagation is also feasible for controlling pollination techniques such as cross pollination, self-pollination or hand pollination in the hybrid breeding program. Moreover, this method of hormone free AR formation and clonal propagation is useful for woody species that are often vegetatively propagated by stem cuttings.

AUTHOR CONTRIBUTIONS

ZC initiated the project. SS and TWC performed experiments. AY contributed to aeroponics plant care. SS, TWC, JH, and ZC wrote the manuscript.

FUNDING

NIE AcRF grant (RI 3/13 CZ) and MOE Tier 1 grant (RP 1/14 CZ).

ACKNOWLEDGMENTS

The project was funded by NIE AcRF grant (RI 3/13 CZ) and MOE Tier 1 grant (RP 1/14 CZ). The authors wish to thank Dr. Kin Wai Lai from Singapore Polytechnic for providing necessary facilities for microtome sectioning in the course of work.

REFERENCES

Abarca, D., and Díaz-Sala, C. (2009). "Adventitious root formation in conifers," in *Adventitious root formation of forest trees and horticultural plants–from genes to applications*, eds K. Niemiand and C. Scagel (Kerala: Research Signpost Publishers), 227–257.

Cheng, F., Liu, S., Wu, J., Fang, L., Sun, S., Liu, B., et al. (2011). BRAD, the genetics and genomics database for Brassica plants. *BMC Plant. Biol.* 11:136. doi: 10.1186/1471-2229-11-136

Douglas, J. S. (1982). *Advanced Guide to Hydroponics*. Dehra Dun: Natraj Publishers.

Ellis, R. (2007). "The Encyclopedia of Seeds," in *Science Technology and Uses*, eds M. Black, J. D. Bewley, and P. Halmer (Wallingford,CT: CABI), 900. ISBN-9780851997230.

Fink, S. (1999). "Root regeneration" in *Pathological and Regenerative Plant Anatomy. Encyclopedia of plant anatomy*, ed. S. Fink (Gebr.Bornträger: Berlin), 6, 350–371.

He, J. (2014). "Elevated root-zone CO2 on photosynthesis of vegetable crops at different air and root-zone temperatures," in *Photosynthesis: Functional Genomics, Physiological Processes and Environmental Issues*, ed. N. Khan (New York, NY: Nova Science Publishers), 23–54.

He, J., Qin, L., and Lee, S. K. (2013). Root-zone CO2 and root-zone temperature effects on photosynthesis and nitrogen metabolism of aeroponically grown lettuce (Lactuca sativa L.) in the tropics. *Photosynthetica* 50, 330–340. doi: 10.1007/s11099-013-0030-5

Legué, V., Rigald, A., and Bhalerao, R. P. (2014). Adventitious root formation in tree species: involvement of transcription factors. *Physiol. Plant.* 151, 192–198. doi: 10.1111/ppl.12197

Luo, H. Y., He, J., and Lee, S. K. (2012). Interaction between potassium (K) concentration and root-zone temperature (RZT) on growth and photosynthesis of temperate lettuce grown in the tropics. *J. Plant Nutr.* 35, 1004–1021. doi: 10.1080/01904167.2012.671404

Parkin, I. A. P., Koh, C., Tang, H. B., Robinson, S. J., Kagale, S., Clarke, W. E., et al. (2014). Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome Biol.* 15, R77. doi: 10.1186/gb-2014-15-6-r77

Rakow, G. (2004). "Species origin and economic importance of Brassica," in *Biotechnology in Agriculture and Forestry,* eds E. C. Pua and C. J. Douglas (New York, NY: Springer-Verlag Berlin Heidelberg), 54, 3–11.

Rochfort, S. J., Imsic, M., Jones, R., Trenerry, V. C., and Tomkins, B. (2006). Characterisation of flavonol conjugates in immature leaves of Pak Choi (*Brassica rapa* L). Ssp chinensis L. (Hanelt) by HPLC-DAD and LC-MS/MS. *J. Agric. Food Chem.* 54, 4855–4860.

Velasco, P., Cartea, M. E., Gonzalez, C., Vilar, M., and Ordas, A. (2007). Factors affecting the glucosinolate content of kale (*Brassica oleracea* acephala group). *J. Agric. Food Chem.* 55, 955–962. doi: 10.1021/jf0624897

Weathers, P. J., and Giles, K. L. (1988). Regeneration of plants using nutrient mist culture. *In Vitro Cell. Dev. Biol. Anim.* 24, 727–732.

Welander, M., Geier, T., Smolka, A., Ahlman, A., Fan, J., and Zhu, L. H. (2014). Origin, timing, and gene expression profile of adventitious rooting in *Arabidopsis* hypocotyls and stems. *Am. J. Bot.* 101, 1–12. doi: 10.3732/ajb.1300258

Wellburn, A. R. (1984). The spectral determination of chlorophylls A and B, as well as total Carotenoids. using various solvents with spectrophotometers of different resolution. *J. Plant Phys.* 144, 307–313. doi: 10.1007/s11120-013-9834-1

Xu, L., and Huang, H. (2014). Genetic and epigenetic controls of plant regeneration. *Curr. Top. Dev. Biol.* 108, 1–33. doi: 10.1016/B978-0-12-391498-9.00009-7

Yang, Q., Chauvin, J. E., and Herve, Y. (1992). A study of factors affecting anther culture of cauliflower (*Brassica oleracea* var. *botrytis). Plant Cell Tiss. Org. Cult.* 28, 289–296. doi: 10.1007/BF00223783

Zhao, J. (2007). *The Genetics of Phytate Content and Morphological Traits in Brassica Rapa*. Ph.D. thesis, Wageningen University, Wageningen.

Zobel, R. W. (1989). "Steady-state control and investigation of root system morphology," in *Steady State and Continuous Control of Root Growth*, eds J. G. Torrey and L. Winship (Amsterdam: Kluwer Academic Publishers), 165–182.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Srikanth, Choong, Yan, He and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

REVIEW
published: 04 September 2015
doi: 10.3389/fpls.2015.00688

Molecular breeding in *Brassica* for salt tolerance: importance of microsatellite (SSR) markers for molecular breeding in *Brassica*

Manu Kumar[1], Ju-Young Choi[1], Nisha Kumari[2], Ashwani Pareek[3] and Seong-Ryong Kim[1]**

[1] Plant Molecular Biology Laboratory, Department of Life Science, Sogang University, Seoul, South Korea, [2] College of Medicine, Seoul National University, Seoul, South Korea, [3] Stress Physiology and Molecular Biology Laboratory, School of Life Science, Jawaharlal Nehru University, New Delhi, India

OPEN ACCESS

Edited by:
Sarvajeet Singh Gill, Maharshi Dayanand University, India

Reviewed by:
Juan Francisco Jimenez Bremont, Instituto Potosino de Investigacion Cientifica y Tecnologica, Mexico
Narottam Dey, Visva-Bharati University, India

***Correspondence:**
Manu Kumar and Seong-Ryong Kim, Plant Molecular Biology Laboratory, Department of Life Science, Sogang University, Seoul 121-742, South Korea
manukumar007@gmail.com; manukumar7@sogang.ac.kr; sungkim@sogang.ac.kr

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *22 April 2015*
Accepted: *20 August 2015*
Published: *04 September 2015*

Citation:
Kumar M, Choi J-Y, Kumari N, Pareek A and Kim S-R (2015) Molecular breeding in Brassica for salt tolerance: importance of microsatellite (SSR) markers for molecular breeding in Brassica. Front. Plant Sci. 6:688. doi: 10.3389/fpls.2015.00688

Salinity is one of the important abiotic factors for any crop management in irrigated as well as rainfed areas, which leads to poor harvests. This yield reduction in salt affected soils can be overcome by improving salt tolerance in crops or by soil reclamation. Salty soils can be reclaimed by leaching the salt or by cultivation of salt tolerance crops. Salt tolerance is a quantitative trait controlled by several genes. Poor knowledge about mechanism of its inheritance makes slow progress in its introgression into target crops. *Brassica* is known to be a good reclamation crop. Inter and intra specific variation within *Brassica* species shows potential of molecular breeding to raise salinity tolerant genotypes. Among the various molecular markers, SSR markers are getting high attention, since they are randomly sparsed, highly variable and show co-dominant inheritance. Furthermore, as sequencing techniques are improving and softwares to find SSR markers are being developed, SSR markers technology is also evolving rapidly. Comparative SSR marker studies targeting *Arabidopsis thaliana* and *Brassica* species which lie in the same family will further aid in studying the salt tolerance related QTLs and subsequent identification of the "candidate genes" and finding out the origin of important QTLs. Although, there are a few reports on molecular breeding for improving salt tolerance using molecular markers in *Brassica* species, usage of SSR markers has a big potential to improve salt tolerance in *Brassica* crops. In order to obtain best harvests, role of SSR marker driven breeding approaches play important role and it has been discussed in this review especially for the introgression of salt tolerance traits in crops.

Keywords: *Brassica*, salt stress, abiotic stress, SSR markers, QTL

Introduction

Salinity in the soil is one of the serious obstacles for agriculture, due to which large areas of the agricultural lands are becoming unfertile. Three fourth of the total Earth surface is covered by saline water and hence significant proportion of this Earth is affected by saline conditions. Over 830 million hectares of land area in the entire earth are salt affected, either by saline water (403 million hectares) or by the conditions related with sodicity (434 million hectares; FAO, 2008) and it is more than six percent of the entire land area in the world. An excess amount of NaCl occurs

as an abiotic environmental factor in many places such as salt deserts in the arid and semi-arid areas, coastal salt marshes and inland saline lakes (Kumar, 2013). During the last decades, apart from the natural salinity, salinization of soils due to intensive agriculture and irrigation has also been becoming a major problem in agriculture. When plants are exposed to salinity, it causes ion imbalance, ion toxicity and hyper osmotic stress (Yamaguchi and Blumwald, 2005; Kumar et al., 2013, 2014). It severely retards the crop growth and productivity. For most of the crops concentrations of 150 mM NaCl are highly toxic though, for a few crops, as low as 25 mM NaCl is lethal. Two main courses of actions were given special importance for providing the solution for salinity stress problem (Epstein, 1985; Ashraf, 1994; Flowers and Yeo, 1995; Grieve et al., 1999), which includes reclamation of saline soils by use of chemicals or by growing salt tolerant plants in the saline soils. Considering its low cost, feasible, and efficient approach, the latter strategy was being emphasized by many plant scientists during the past few decades. This includes cereals, legumes and other commercially important crops.

Apart from the cereals and legumes, oil seeds are very important for human food and are at the third position among the crops. At least forty different plant species are known to be grown for the oils production (Weiss, 1983). Among the oilseed crops, *Brassicas* which belongs to the family *Brassicaceae* are very important oilseed crops. The family *Brassicaceae* includes various crops, which are rich in nutritional and economic values. The members of the *Brassica* genus are sometimes collectively called as cabbages/mustards/-cole crops. *Brassica* contains more number of important horticultural and agricultural crops. The members of the *Brassica* genus also contains more number of weed species and wild relatives, making it a perfect platform for crop improvement practices, due to the presence of wide genetic base. Apart from the oilseeds (mustard seed, oilseed rape), almost every part of the plant of some species or the other are edible and grown for food, which includes the stems (kohlrali), root (Swedes, turnips), flower (Cualiflower, broccoli), and leaves (Cabbages, Brussels sprout). *Brassica* vegetables are most commonly regarded for their nutritional and medicinal properties (Beecher, 1994; Carvalhoa et al., 2006). They contain high amounts of soluble fiber and vitamin C (Divisi et al., 2006). *Brassica* contains different nutrients having potential anticancer properties like 3, 3′ diindolylmethane, sulforaphane, and selenium (Finley et al., 2005; Banerjee et al., 2012). Since *Brassicas* are of high agricultural importance, they are of much scientific interest.

Before discussing the markers and salinity tolerance in *Brassica*, it is important to know the relationship with in the *Brassica* species.

Relationships between Crop *Brassicas*

The relationship between six particular species in *Brassica* genus (*B. carinata*, *B. juncea*, *B. napus*, *B. nigra*, *B. oleraceae,* and *B. rapa*) is well described by the Triangle of U theory (Nagaharu, 1935; **Figure 1**). The Triangle of U theory explains the high chromosome number species [*B. carinata* (BBCC), $2n = 34$; *B. juncea* (AABB), $2n = 36$; and *B. napus* (AACC), $2n = 38$] which are amphidiploids and possibly formed through the interspecific hybridization between the low chromosome number species in pairs [*B. nigra* (BB), $2n = 16$; *B. oleraceae* (CC), $2n = 18$; and *B. rapa* (AA), $2n = 20$].

Salinity Tolerance in *Brassica*

Salt stress tolerance is one of the highly complex processes in number of plant species. Many complex mechanisms are involved at different plant developmental levels. But some of these mechanisms are functional at a particular time in a given species. Apart from this, the effect of one process can exclude the effect of the other process at a particular time (Gorham et al., 1991; Ashraf, 1994; Yeo, 1998; Carvajal et al., 1999). Salt stress tolerance in plants is a developmentally regulated phenomenon and tolerance at one stage of development sometimes may not correlate with the stress tolerance at other stages. For example, in barely, corn, rice, tomato, and wheat, salt stress tolerance tends to increase as the plants become older. The situation becomes more complicated with polyploidy species in comparison with their respective diploid parents. Polyploid species can withstand the harmful environmental factors such as salt stress tolerance better than their own respective diploid parents. Different studies gives the indications that the amphidiploid species, *Brassica carinata*, *B. juncea*, and *B. napus* have the superiority over the diploid species, *B. campestries*, *B. nigra,* and *B. oleraceae* in terms of salinity tolerance (Ashraf et al., 2001). It was also found that the amphidiploid genotypes, *B. carinata* and *B. napus* were salt stress tolerant when compared with *B. campestries*. Another amphidiploid *B. juncea* is intermediate in salt stress tolerance (Ashraf and McNeilly, 1990). The continued survival of salt stress tolerant plants and the differences between the genotypes with salt sensitive plant species point out the presence of a genetic basis of salt stress tolerance.

Utilization of *Brassica* Genetic Diversity

In the presence of environmental stresses, such as drought, salt, cold, nutrient deficiency, and water logging, growth of the *Brassica* plants, their oil production and reproduction capabilities are reduced. Hence, these *Brassica* species are normally grown on standard non-saline conditions in order to maximize the yield. If they are grown on salt affected soils, yield losses are severe. Therefore, improvement of their salt stress tolerance is of considerable economic value. For breeding program to be successful, presence of a significant heritable distinction with in the gene pool of these crops is a compulsory requirement (Becker et al., 1995; O'Neill et al., 2003; Kumar, 2015). Because of the close relationship and the presences of important inter and intra specific distinctions within Brassica species, the breeding programs for salinity stress tolerance have been highly benefited. It is also believed that support from other approaches such as mutagenesis, fusion of protoplast or recombinant technology can

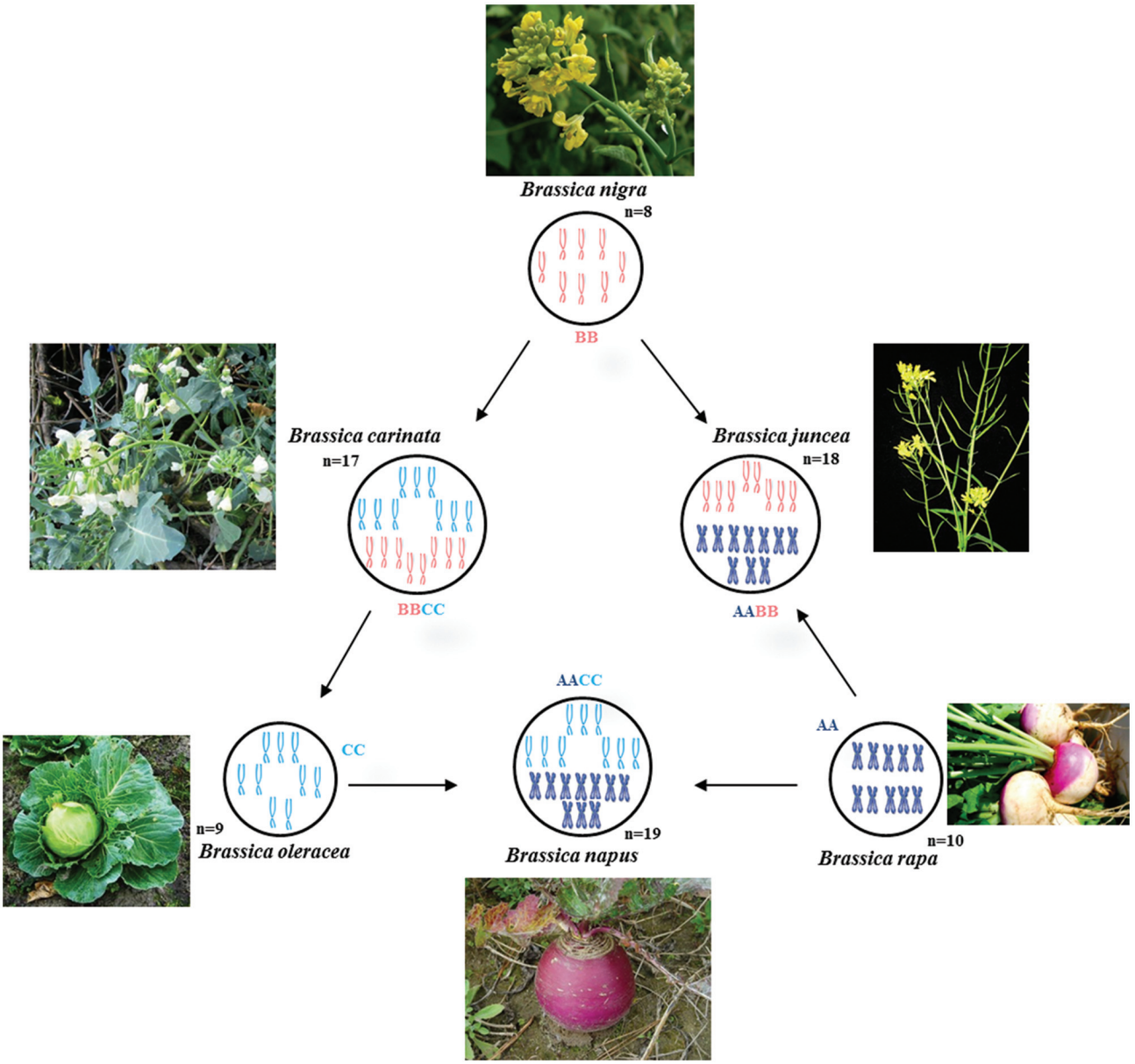


FIGURE 1 | Triangle of U diagram illustrating the genetic relationships between six species of the genus *Brassica*, such as Chinese cabbage (AA), Cabbage (CC), Rapeseed (AACC), Indian mustard (AABB), and other common vegetables, showing the genetic relationships between Chromosomes from each of the genomes A, B, and C are represented by different colors.

also be helpful in achieving the desired target (Diers et al., 1996; Riaz et al., 2001; Seyis et al., 2003). Even though there is a great inter and intra specific variation for salt stress tolerance within *Brassica* species, generation of new variation through induced mutation and utilization of those new variants gives more scope for enhancing salt stress tolerance. Due to the advancements of molecular techniques, the mutants can be identified and analyzed using DNA fingerprinting and mapping on PCR based markers such as SSR, RAPD, AFLP, and STMS (Diers and Osborn, 1994; Halldén et al., 1994; Thormann et al., 1994; Plieske and Struss, 2001a).

Molecular Markers in Breeding for Salinity Tolerance

Molecular genetics is one of the most important technologies in the today's world. Stress tolerance and yield are difficult to breed using conventional methods because of their polygenic nature and are also largely influenced by environment and genotype. The complex quantitative feature of the most mechanisms involved in the salt stress tolerance is the main reason for the limited success of the modern salt tolerance breeding approaches (Yeo and Flowers, 1986). The association and application of the

indirect selection markers which are genetically linked with the trait (s) of interest is a well-known approach for the betterment of the crop having difficult traits which includes salt stress tolerance (Im et al., 2014a). DNA marker technology has revolutionized the genome research and breeding in the recent decades. Implementation of various available markers and QTL mapping techniques have contributed for the good knowledge of genetic bases of various agriculturally significant traits such as resistance to biotic stresses, abiotic stresses tolerance, yield and nutritional quality in various crops (Xue et al., 2010; Ali et al., 2013).

Since, breeders use QTL linked markers to find the position of markers on the loci that controls the concerned traits; the number of methods to identify the phenotype is reduced. Therefore, the necessity for large-scale methods over time and space is significantly reduced. A few salt stress related QTL detected by SSR markers have been listed in the **Table 1**.

In hexaploid bread wheat (*Triticum aestivum*), important locus (*Kna1*) has been reported that regulate the transport of Na^+/K^+ from root to shoot specifically, by containing a lower Na^+/K^+ ratio within the leaves (Gorham et al., 1987, 1990; Dubcovsky et al., 1996; Luo et al., 1996). Meanwhile, in durum *Triticum turgidum* L. ssp. durum Desf. (wheat) discharge process of Na^+ is linked to *Nax1* (Na^+ exclusion 1; Huang et al., 2006, 2008), that might be related to the HKT8 (HKT1;5) and HKT7 (Na^+ transporters HKT1;4). It has been reported that *Nax1* loci efficiently reduce Na^+ passage to shoot from root, by keeping Na^+/K^+ balance within the leaf of wheat by loading K^+ into and excluding Na^+ from, the xylem of the plant (James et al., 2006). Using F2 population of a hybrid within indica rice cultivar 'IR36' and japonica rice cultivar 'Jiucaiqing,' two QTLs identified for root Na^+/K^+ ratio, which were mapped to chromosomes 2 and 6 (Yao et al., 2005). For Salt tolerance traits different QTLs have been recognized in rice which include those at chromosome number 1- *Saltol* QTL, *QNa,* and *SKC1/OsHKT8* along with, QNa:K on chromosome 4. *Saltol* describes many changes for the uptake of ion during salinity stress (Bonilla et al., 2002; Gregorio et al., 2002). For highest Na^+ uptake *QNa* is QTL (Flowers et al., 2000). For Na^+/K^+, QNa:K is the corresponding QTL (Singh et al., 2001). For regulation of K^+/Na^+ ratio for homoeostasis in salt stress tolerant indica cultivar 'Nona Bokra' *SKC1/OsHKT8* is the corresponding QTL (Lin et al., 2004; Ren et al., 2005). Also, many other QTLs are on the every chromosome except chromosome nine in the root for Na^+/K^+ ratio, and for exchange of ion three QTLs on chromosomes 10 and 3 (Sabouri and Sabouri, 2008), for tissue Na^+/K^+ ratio four QTLs and each for Na^+ and K^+ uptake on various chromosomes one QTL (Lang et al., 2001). Thereafter, 14 QTLs was identified for shoot and root Na^+/K^+ ratio and Na^+ and K^+ content on different rice chromosomes, recently (Ahmadi and Fotokian, 2011). Among these QTLs, on chromosome 1 for root K^+ content, *QKr1.2* was identified as one of very bright QTL as it explained around 30% of the variation of observed salt stress tolerance in rice. Furthermore,

TABLE 1 | Identification of quantitative trait loci (QTLs) by SSR markers for salt tolerance (ST) in different plant species.

Crop plants	Locus	Traits governed	Reference
Wheat (*Triticum aestivum* L.)	*Kna1*	Controls the selectivity of Na^+ and K^+ transport from root to shoot and maintains high K^+/Na^+ ratio	Gorham et al. (1990), Dubcovsky et al. (1996)
	Nax1	Both are involved in decreasing Na^+ uptake and enhancing K^+ loading into the xylem	Lindsay et al. (2004), Huang et al. (2006)
Rice (*Oryza sativa* L.)	*qRL-7, qDWRO-9a* and *qDWRO-9b qBI-1a* and *qBI-1b*	Play important roles in root length and root dry weight at seedling stage under saline conditions	Sabouri and Sabouri (2008)
	QNa, QNa:K, SKC1/OsHKT8	Regulate K^+/Na^+ homoeostasis	Ren et al. (2005)
	qDM-3 and *qDM-8, qSTR-6*	Improve Na^+/K^+ ratio under saline conditions	Sabouri (2009)
	qNAK-2 and *qNAK-6*	Improve Na^+/K^+ ratio	Yao et al. (2005)
	Saltol	Controls shoot Na^+/K^+ homoeostasis	Thomson et al. (2010)
	Saltol and *non-Saltol*	Control shoot Na^+/K^+ homoeostasis	Alam et al. (2011)
	QKr1.2	Controls K^+ content in root	Ahmadi and Fotokian (2011)
Barley (*Hordeum vulgare*)	Five QTL for ST were identified on chromosomes 1H, 2H, 5H, 6H, and 7H, which accounted for more than 50% of the phenotypic variation	Enhance vegetative growth under saline stress	Zhou et al. (2012)
	A locus *HvNax3* on the short arm of chromosome 7H in wild barley (*Hordeum vulgare* ssp. *spontaneum*) accession CPI-71284-48	Reduces shoot Na^+ content by 10–25% in plants grown under salt stress (150 mM NaCl)	Shavrukov et al. (2010)
White clover (*Trifolium repens* L.)	Several QTLs for ST, some at common locations, but each of low scale	Affect ST during vegetative stage	Wang et al. (2010)
Soybean (*Glycine max* (L.) Merr.)	A major QTL for ST was identified near the Sat091 SSR marker on linkage group (LG) N	Maintains growth under salt stress	Lee et al. (2004)
	Eight QTLs for ST were detected	Maintains growth under salt stress	Chen et al. (2008)
	A major QTL for ST was detected	Maintains healthy growth under salt stress	Hamwieh et al. (2011)

on rice chromosomes 8 and 10, two newly identified QTLs (*SalTol*8-1 and *SalTol*10-1) based on an F2 hybrid of a cross between a high salt stress tolerant line (IR61920-3B-22-2-1) and a medium salt stress tolerant line (BRRI-dhan40; Islam et al., 2011).

Also, in *Hordeum vulgare* L. (barley, **Table 1**), many studies have discovered QTLs for salt stress tolerance related phenotypes. Recently, 30 QTLs were identified for 10 different traits, such as K^+ and shoot Na^+ content, yield-related traits, several growth and Na^+/K^+ ratio, in populations grown on normal soil and salt affected soil. In the three species of *Helianthus* sp. (sunflower) and *Helianthus paradoxus*, ion-uptake traits related QTL analysis from highly salt affected habitat and its relative ancestor *H. petiolaris* and *H. annuus* which are both relatively salt sensitive, identified 14 ion uptake QTLs (Lexer et al., 2003). Additional studies are required to decide the benefits of unreported QTLs within crop breeding to improve salt stress tolerance. Since molecular marker techniques for breeding is economical and rapid, this technique is a very powerful method to enhance breeding programs to improve plant tolerance toward salinity. Especially, DNA markers are very important in plant breeding for the selection of polygenic traits, because of the absence of genotype X environment interaction, epistatic effect, and also ease in the picking up of homozygous plants and the homozygous lines can be greatly distinguished from the others at an early generation. Before the crosses of parental lines, molecular characterization of germplasm can help the genetic variations among the parental genotypes increase. Genetic diversity present in the breeding population is maximized and the labor time that is required for either direct selection in traditional breeding or in direct selection through QTLs minimizes. Even though this kind of procedure remains encouraging, its implementation to the complex traits such as salt stress tolerance may be limited due to the close genetic relationship between, wide confidence intervals and, big sample size requirement for screening of the segregating populations, parental population, and possible interactions between genotype and environment for QTL study.

Available Marker Systems in *Brassica*

In *Brassica*, genome research with the application of marker assisted program began to emerge in the late 1980s when the first RFLP linkage map for *B. oleraceae* (Slocum et al., 1990), *B. napus* (Landry et al., 1991), and *B. rapa* (Song et al., 1991) was developed. For phylogenetic studies and genetic mapping in *Brassica*, RFLPs and RAPDs have been extensively used (Williams et al., 1990). However, the discovery of the PCR (Mullis and Faloona, 1987) leads the potential to increase the variety and density of marker in the already existing genetic maps with ISSR, AFLP and with the microsatellites (Grist et al., 1993), also called as SSR. SSRs are highly important resource of map-based alignment among distinct crosses, because of their robust, simple, and relatively inexpensive analysis and highly polymorphic nature. The number of available *Brassica* SSRs (microsatellite) primers is increasing (http://www.brassica.info/ssr/SSRinfo.htm) the list is given in the **Table 2**. *Brassica* genome integration greatly assisted the release of highly polymorphic mapped based, robust SSR markers of the entire *B. nigra, B. rapa*, *B. napus,* and *B. oleracea* genome into public domain. A large number of SSRs (microsatellite) markers have been developed among the cultivated *Brassica* species such as *B. oleracea* (AACC) and the diploids *B. rapa* (AA), *B. nigra* (BB), and *B. oleracea* (CC) which have been shown to be applicable within and between different *Brassica* species. One of the main limit to develop SSR markers in some *Brassica* crops is the lack of finished genome sequence. However, thanks to development of sequencing technology, *B. rapa, B. oleracea,* and *B. napus* are sequenced, recently (Wang et al., 2011; Yu et al., 2013; Shi et al., 2014; Sharma et al., 2015). From this sequence, 140998, 229389, 420991 mono- to hexanucleotide repeat microsatellites are identified using PERL5 script MIcroSAtellite (Thiel et al., 2003). From these identified microsatellites, 115869, 185662, and 356522 SSR markers were developed using *in silico* method, respectively (Shi et al., 2014; **Table 3**). In the past few years, the research work has clearly proven the power of candidate gene studies and genetic maps of high density for the location of molecular markers that are closely linked with the useful trait (s) within *Brassica* have been developed and most of them have been successfully integrated into the *Brassica* oilseed breeding programs (**Table 4**).

TABLE 2 | The number of available *Brassica* microsatellite primers in public domain.

Microsatellite set	No. available in public domain
BBRC microsatellite program	397
HRI set	6
Kresovich and Szewc-McFadden	24
Lagercrantz et al. (1993)	5
AAFC Consortium (commercial)	80
Suwabe et al. (2002, 2003)	38
Uzunova and Ecke (1999)	8
Bell and Ecker (1994; *Arabidopsis* microsatellites)	30
INRA Versailles (*Arabidopsis* microsatellites)	120
Celera Consortium (commercial)	171
Oilcrops (http://oilcrops.info//SSRdb)	3974
Total	4853

Importance of Microsatellite (SSR) Marker

Over the past few years, various new PCR based marker such as AFLPs, RAPDs, and microsatellites have been developed and applied in crop improvement program. Microsatellites markers have great deal of potential among all the markers. Litt and Luty (1989) coined the term microsatellite. Microsatellites markers are also defined as simple sequence repeats (SSRs) which are based on unique DNA sequences that are flanking short repetitive traits of simple sequence motifs, for example – di or tri nucleotides. They are randomly distributed within the eukaryotic genomes (Smith and Devey, 1994). They are variable with respect to the number of repeats, pedigree analysis and are highly efficient in

TABLE 3 | Microsatellite sequence from the genome of *Brassica rapa, B. oleracea, B. napus* (Shi et al., 2014), and *Arabidopsis thaliana* (*Arabidopsis* Genome Initiative, 2000).

Motif	*B. napus*		*B. oleracea*		*B. rapa*		*A. thaliana*	
	Minimum repeat number	Number (%)	Minimum repeat number	Number (%)	Minimum repeat number	Number (%)	Minimum repeat number	Number (%)
Mono	12	97 128 (23.1)	12	55 433 (24.2)	12	31 258 (22.2)	12	13 650 (43.4)
A	12	94 281 (22.4)	12	52 021 (22.7)	12	29 536 (20.9)	12	13 434 (42.7)
C	12	2847 (0.7)	12	3412 (1.5)	12	1722 (1.2)	12	216 (0.7)
Di	6	98 816 (23.5)	6	55 336 (24.1)	6	33 885 (24.0)	6	7433 (23.6)
AT	6	57 070 (13.6)	6	33 315 (14.5)	6	19 697 (14.0)	6	852 (2.7)
AG	6	34 638 (8.2)	6	18 593 (8.1)	6	11 683 (8.3)	6	2034 (6.5)
AC	6	7072 (1.7)	6	3411 (1.5)	6	2490 (1.8)	6	4544 (14.4)
CG	6	36 (0.0)	6	17 (0.0)	6	15 (0.0)	6	3 (0)
Tri	4	91 448 (21.7)	4	47 716 (20.8)	4	32 387 (23.0)	4	9479 (30.1)
AAG	4	29 395 (7.0)	4	15 322 (6.7)	4	9796 (6.9)	4	1435 (4.6)
AAT	4	17 722 (4.2)	4	9355 (4.1)	4	6334 (4.5)	4	3304 (10.5)
ATC	4	12 314 (2.9)	4	6093 (2.7)	4	4211 (3.0)	4	717 (2.3)
AAC	4	10 046 (2.4)	4	5036 (2.2)	4	3637 (2.6)	4	611(1.9)
AGG	4	8349 (2.0)	4	5425 (2.4)	4	3243 (2.3)	4	411 (1.3)
ACC	4	5969 (1.4)	4	2788 (1.2)	4	2144 (1.5)	4	120 (0.4)
AGC	4	2878 (0.7)	4	1390 (0.6)	4	1127 (0.8)	4	509(1.6)
ACT	4	1754 (0.4)	4	818 (0.4)	4	674 (0.5)	4	677 (2.2)
CCG	4	1566 (0.4)	4	799 (0.3)	4	626 (0.4)	4	1633 (5.2)
ACG	4	1455 (0.3)	4	690 (0.3)	4	595 (0.4)	4	62 (0.2)
Tetra	3	91 268 (21.7)	3	48 394 (21.1)	3	29 433 (20.9)	4	462 (1.5)
AAAT	3	33 903 (8.1)	3	18 856 (8.2)	3	11 870 (8.4)	4	98 (0.3)
AAAG	3	10 795 (2.6)	3	5697 (2.5)	3	3496 (2.5)	4	94 (0.3)
AAAC	3	9717 (2.3)	3	4723 (2.1)	3	3333 (2.4)	4	73 (0.2)
AATT	3	7863 (1.9)	3	4088 (1.8)	3	2534 (1.8)	4	24 (0.1)
AATC	3	5112 (1.2)	3	2956 (1.3)	3	1137 (0.8)	4	47 (0.1)
Others	3	23 878 (5.7)	3	12 074 (5.3)	3	7063 (5.0)	4	126(0.4)
Penta	3	29 058 (6.9)	3	15 012 (6.5)	3	9856 (7.0)	4	283 (0.9)
AAAAT	3	7617 (1.8)	3	4051 (1.8)	3	2758 (2.0)	4	14 (0)
AACCG	3	3541 (0.8)	3	2169 (0.9)	3	1000 (0.7)	4	8 (0)
AAAAC	3	2424 (0.6)	3	1186 (0.5)	3	878 (0.6)	4	34 (0.1)
AAAAG	3	2351 (0.6)	3	1164 (0.5)	3	683 (0.5)	4	25 (0.1)
AAATT	3	1488 (0.4)	3	784 (0.3)	3	568 (0.4)	4	2 (0)
AAACC	3	1273 (0.3)	3	713 (0.3)	3	490 (0.3)	4	10 (0)
AATAT	3	1083 (0.3)	3	573 (0.2)	3	394 (0.3)	4	1 (0)
Others	3	9281 (2.2)	3	4372 (1.9)	3	3085 (2.2)	4	189 (0.6)
Hexa	3	13 273 (3.2)	3	7498 (3.3)	3	4179 (3.0)	4	149 (0.5)
AAAAAT	3	1803 (0.4)	3	987 (0.4)	3	628 (0.4)	4	2 (0)
AAAATT	3	1341 (0.3)	3	776 (0.3)	3	267 (0.2)	5	1 (0)
AAAAAC	3	746 (0.2)	3	395 (0.2)	3	269 (0.2)	4	7 (0)
AAAAAG	3	644 (0.2)	3	357 (0.2)	3	175 (0.1)	4	4 (0)
AAATAT	3	436 (0.1)	3	246 (0.1)	3	150 (0.1)	-	0 (0)
Others	3	8303 (2.0)	3	4737 (2.1)	3	2690 (1.9)	4	135 (0.4)
Total	3	420991 (100)	3	229389 (100)	3	140998 (100)	4	31456 (100)

Arabidopsis thaliana microsatellite sequence was found from using msatcommander software (Faircloth, 2008). This table only shows particular sequence motifs data with whole motifs up to hexanucleotide is in Supplementary Table S1.

the fingerprinting and show co-dominant inheritance of different crops. The extra ordinary level of instructive polymorphism at SSR locus originated from the apparent tendency of development or replication or unequal crossing-over event at the time of meiosis. The strengths of SSR markers include their high numbers in eukaryotes, the codominance of alleles, and their arbitrary dispensation throughout the genome with special consortium with in low-copy regions (Morgante et al., 2002).

TABLE 4 | Study of important traits in *Brassica* species.

Character	Trait	Species	Reference
Morphological traits	Seed coat color	*B. napus*	Somers et al. (2001)
"	"	*B. juncea*	Padmaja et al. (2005)
"	"	*B. rapa*	Mukhlesur et al. (2007)
	Seed coat color	*B. napus*	Rezaeizad et al. (2011)
	Seed coat color	*B. napus*	Qu et al. (2015)
	Seed coat color	*B. napus*	Gustafson et al. (2006), Fu et al. (2007), Xiao et al. (2007), Francki et al. (2010)
"	Yield influencing QTLs	*B. juncea*	Ramchiary et al. (2007)
"	Flowering time	*B. napus*	Ferreira et al. (1995b)
"	Root morphology	*B. rapa*	Lu et al. (2008)
"	Plant height	*B. napus*	Foisset et al. (1995), Barret et al. (1998)
"	Petal – less flower	*B. napus*	Fray et al. (1997)
Oil content	Glucosinolate content	*B. napus*	Uzunova et al. (1995)
"	Oleic acid	*B. napus*	Hu et al. (1999)
"	Linoleic acid	*B. napus*	Somers et al. (1998), Hu et al. (2006)
"	Linolenic acid	*B. napus*	Hu et al. (1995), Tanhuanpää and Schulman (2002)
"	Erucic acid	*B. napus*	Ecke et al. (1995)
"	"	*B. juncea*	Gupta et al. (2004)
"	Seed glucosinulates	*B. napus*	Uzunova et al. (1995), Howell et al. (2003)
	Oil content	*B. napus*	Qu et al. (2015)
	Fiber content	*B. napus*	Gustafson et al. (2006), Fu et al. (2007), Xiao et al. (2007), Francki et al. (2010)
	Glucosinolate content	*B. napus*	Hasan et al. (2008)
Disease resistance	*Albugo candida*	*B. napus*	Ferreira et al. (1995c), Kole et al. (2002b)
"	*Plasmodiophora brassicae*	*B. pekinensis*	Lee et al. (2002)
"	*Leptosphaeria maculans*	*B. napus*	Dion et al. (1995), Ferreira et al. (1995a), Pilet et al. (1998a,b, 2001)
"	Turnip mosaic virus	*B. napus*	Walsh et al. (1999), Dreyer et al. (2001)
"	*Sclerotiana sclerotiorum*	*B. napus*	Zhao and Meng (2003)
"	*Phoma*	*B. napus*	Plieske and Struss (2001b)
"	*Pyrenopeziza brassicae*	*B. napus*	Pilet et al. (1998b)
Abiotic stress	Winter survival	*B. napus*	Kole et al. (2002a), Asghari et al. (2007)
	Drought and cold	*B. napus*	Kagale et al. (2007)
Male sterility	'Ogura' fertility restorer	*B. napus*	Delourme et al. (1998), Brown et al. (2003)
"	'Polima' fertility restorer	*B. napus*	Jean et al. (1998)
"	'Kosena' fertility restorer	*B. napus*	Imai et al. (2003)
"	*tournefortii* fertility restorer	*B. napus*	Janeja et al. (2003)

Also, low quantities of template DNA (10–50 ng/reaction) are required, because of the PCR based technique. Reproducibility of SSR markers is high due to the use of lengthy PCR primers (Im et al., 2014b), and its use does not even require high quality DNA. Even though, the RFLP was one of the first markers used for genome analysis, RFLP technique is laborious and RFLP is less polymorphic than SSR marker. Also, improved technique that is more simple and efficient to find polymorphism in SSR marker makes SSR marker more useful (Kumar et al., 2015). Furthermore, since conventional microsatellite generating method using genomic libraries (Weising et al., 2005; Smykal et al., 2006) was replaced by *in silico* microsatellite generating method, many software generating microsatellite was made, such as MISA, MicrosatDesign, msatminer, msatcommander, IMEx, WebSat (Thiel et al., 2003; Singan and Colbourne, 2005; Thurston and Field, 2005; Mudunuri and Nagarajaram, 2007; Faircloth, 2008; Martins et al., 2009). These programs and genome sequence information make SSR marker generation procedure convenient. Therefore, the number of available SSR markers increases rapidly and available SSR marker in the genome becomes more dense. Therefore, the amount of SSR marker database is increasing. Even though, there are not so many reports about QTL for salt stress tolerance, many QTL has been identified and QTL information was enhanced by association mapping using SSR markers in

Brassica crops. In *Brassica napus*, 53 SSR markers were found to be significantly associated to three phenolic fractions and 11 markers found to be associated to total phenolic acid contents. Among these markers, four SSR markers are derived from QTL for seed color (Rezaeizad et al., 2011). Twenty five and 11 SSR markers were found to be associated with seed coat color and oil content. Among these markers six SSR markers are associated with both of coat color and oil content (Qu et al., 2015). Also, association study between known major QTL and SSR marker is useful to find candidate gene because of high density of SSR marker in genome. Main QTL for seed color and fiber content on one of the homoeolog chromosomes A9 or C8 in *B. napus* has been described by many studies (Gustafson et al., 2006; Fu et al., 2007; Xiao et al., 2007). SSR markers bridge the sequence contig overlaying this QTL was identified and four of these SSR markers from small genomic region less than 50 kb which are strongly associated with seed color and fiber content traits were identified (Francki et al., 2010). These SSR markers in known major QTL or near these QTL which are strongly associated with the trait are useful to find candidate gene. SSR markers can be analyzed interspecifically. After mapping SSR markers in several species, comparative mapping of SSR markers can be made using alignment of orthologous loci (Yu et al., 2004; Zhang et al., 2007; Yu and Li, 2008). This will help in identifying the origin of QTL, find candidate gene as well as structural collinearity in genome between two species can also be dissected out.

The Candidate Gene Approach

Many traits of agricultural importance, which includes salinity tolerant traits, exhibits quantitative inheritance, which is mostly the result of multiple genes influenced by the environment. Due to the their imprecise localization on the genetic maps and multiplicity of genes defining a quantitative trait and their incomplete effects on phenotypic differentiation, the candidate gene approach is more adapted to the QTL characterization. The candidate gene approach has been emphasized as an encouraging method for combining QTL analysis with the large-scale data available on the cloning and genes characterization (Pflieger et al., 2001). Genes likely to be involved in the biochemical pathways, in this technique, that lead to a trait articulation, are engaged as molecular markers for QTL analysis. The *Brassica* species are the closest crop relative of the model crucifer *Arabidopsis thaliana* and the complete sequencing of this model crop has also created the way to relative examination into the complicated structure of *Brassica* genomes (Chalhoub et al., 2003). The comparative studies of flanked genome regions of known genes shows the extensive co-linearity between *Arabidopsis* and *Brassica* genome segments on a small scale level (Fourmann et al., 1998; Babula et al., 2003). Over the long chromosome stretches, the large scale synteny makes way to use the sequence data from the markers bound QTLs or genes of interest in *Brassica* to determine candidate genes from the chromosome segment in *Arabidopsis*. For example, different homeologs regions in *B. napus* and *B. rapa* that have different QTLs that regulate flowering time, each show useful similarity to the *Arabidopsis* chromosomal part containing a particular number of genes that influence the flowering time (Ferreira et al., 1995b; Lagercrantz et al., 1996).

Furthermore, comparative mapping between *Brassica* and *Arabidopsis* with SSR marker is helpful to identify candidate gene. Comparative mapping based on SSR marker between *B. rapa* and *A. thaliana* shows corresponding regions in *A. thaliana* for *Crr1* and *Crr2* which are QTL for club root resistance in *B. rapa* are in a small region of *A. thaliana* chromosome four where one of the region of disease resistance gene cluster has been identified. Therefore, it seems that the gene for club root resistance in *B. rapa* is related with disease resistance gene cluster in *A. thaliana* (Suwabe et al., 2006).

While comparing the salt stress tolerance of a particular *Brassica* species at the early growth stages, *B. carinata*, *B. juncea,* and *B. napus*, had better salinity tolerance than *B. campestris* (Ashraf and McNeilly, 1990). The reactions of four *Brassica* species, *B. campestris*, *B. carinata*, *B. juncea*, and *B. napus* to four different salts, $CaCl_2$, $MgCl_2$, NaCl, and Na_2SO_4, was tested at the seedling and germination stage using sand culture and solution (Ashraf et al., 1989). Effect of NaCl was more significantly within the effect of four salts and it inhibits the germination rate of all four species. There was no uniform connection between results for seedling growth and germination rate, exceptions in *B. napus*, which showed more seedling growth and better germination rate under the salinity stress as compared to the other three species.

Huge and still growing *Arabidopsis* EST database and the amalgamation into the comparative *Brassica* genome study helps fine mapping of the genomic re-arrangements and the recognition of regions that contains genes crop plants. Also recognition of the association of given traits in *Brassica* crops with the candidate genes of *Arabidopsis* and the generation of the molecular markers that have association with the corresponding genes (Panjabi et al., 2008). For example the Co (constans) gene isolated from *Arabidopsis*, which is involved in late flowering is a putative candidate gene (genes isolated in model species establish the putative CGs for the agronomic species) for two QTLs, which control the flowering time in *B. napus* (Putterill et al., 1995; Robert et al., 1998). Putative CGs involved in fatty acid metabolism, were mapped on the rapeseed genome. In *B. rapa*, a important QTL for flowering time was found in the region homologous to the *Arabidopsis* chromosome five top, where many genes that regulate flowering are located (Kole et al., 2001). In *B. napus*, many cold and drought induced genes were isolated and characterized and there is a huge correlation between the development of freezing tolerance and the expression of some of these genes, which seems to be up-regulated by cold stress (Kole et al., 2002a; Asghari et al., 2007; Kagale et al., 2007).

In *B. oleracea*, fifteen QTLs regulating the flowering time were situated to the *Arabidopsis* genomic segments that contain flowering time genes that affect flowering (Okazaki et al., 2007). Similarly large sets of *Arabidopsis* QTLs were found for *Brassica* QTLs that were influencing leaves and whole plant structure (Lan and Paterson, 2001). Also, based on synteny, SSR markers near the QTL for glucosinolate content in *B. napus* whose orthologs in *A.thaliana* is linked to candidate genes were identified and four putative candidate genes for glucosinolate biosynthesis were identified (Hasan et al., 2008.). Furthermore,

Using *in silico* analysis about *B. oleracea* genome and synteny with *A. thaliana*, putative seven candidate major loci for regulating glucosinolate content were proposed (Sotelo et al., 2014.). Due to this *Arabidopsis* genome sequence is a very informative resource for identifying and further assessment of candidate genes that may account for the control of complicated traits in *Brassica* at the genetic level. But main difficulty in the application of genetic information of *Arabidopsis* for the map-based cloning, candidate gene identification and marker development in *Brassica* crop species is hindered by the complicated structure of the polyploidy *Brassica* genome.

Conclusion

Today's agriculture certainly requires salt tolerant *Brassicas* for the very commercial purpose of the crop. Recent in-depth investigations at physiological and moleuclar levels have identified many ways by which wild type plants cope with salinity stress. Thanks to close relationship and the significant inter and intra specific variation within *Brassica* species which shows huge potential for breeding for salt stress tolerance in *Brassica* crops. Nonetheless, it is clear that to link the salt tolerance trait with QTL position on the chromosome, a proper breeding program assisted with the markers is a prerequisite.

Among the various molecular markers available for this purpose, SSRs are gaining huge attention. SSR markers have many advantages such as high polymorphism, relatively simple methods for identification and most importantly, small amounts of plant material is required for the PCR-based experiment. Other advantages of SSR markars are their random distribution, moderate genome coverage and co-dominant inheritance.

Since *Brassica* and *Arabidopsis* belongs to the same family, many studies and database of *Arabidopsis* are helpful to do breeding in order to improve salt tolerance. Although genome of B. Rapa (485 Mb) and B. Oleracea (630 Mb) is bigger than the genome of A. *thaliana* (Wang et al., 2011; Yu et al., 2013) and *A. thaliana* genome segments are dispersed and rearranged in many *Brassica* crops (Kowalski et al., 1994; Lagercrantz, 1998; Parkin et al., 2005), it has many genomic segments which are highly conserved. Therefore, comparative molecular marker mapping is quite informative. There are 140998, 229389, 420991 mono to hexanucleotide repeat SSRs from the genome of *B. rapa, B. oleracea,* and *B. napus*, respectively, out of which 31456 are microsatellite sequence which are candidates of SSR markers from the genome of *A. thaliana* (**Table 3**, Supplementary Data S1). There is tendancy that major motif sequences of candidates of SSR marker in both *Brassica* crops and *A. thaliana* are A/T rich and the motif sequences that are abundant in the genome of *Brassica* crops are also abundant in *A. thaliana*. Even microsatellite sequence in *A. thaliana* genome is much less than that in *Brassica* crops, the number of SSR marker is enough to show potential of comparative study of salt tolerance QTL using SSR marker in *Brassica* crops and *A. thaliana*. Furthermore, since EST-SSR marker developed from EST-database is more useful to find the candidate genes, the available huge EST-database is helpful for identifying the candidate gene from the QTL. Also, other physiological and agronomical traits can be studied with SSR markers to make a robust and healthy plants with high yield.

Acknowledgments

I thank Mrs. Nisha Kumari for proof reading and editing of this paper. I would also like to thanks Jawaharlal Nehru University New Delhi, and Sogang University Seoul for providing logistics support and space to work. This work was supported by a grant from IAEA, Vienna and Sogang University, Republic of Korea.

Supplementary Material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2015.00688

References

Ahmadi, J., and Fotokian, M. H. (2011). Identification and mapping of quantitative trait loci associated with salinity tolerance in rice (*Oryza sativa*) using SSR markers. *Iran. J. Biotechnol.* 9, 21–30.

Alam, R., Sazzadur, R. M., Seraj, Z. I., Thomson, M. J., Ismail, A. M., Tumimbang-Raiz, E., et al. (2011). Investigation of seedling-stage salinity tolerance QTLs using backcross lines derived from *Oryza sativa* L. Pokkali. *Plant Breeding* 130, 430–437. doi: 10.1111/j.1439-0523.2010.01837.x

Ali, A., Yang, E. M., Lee, S. Y., and Chung, S. M. (2013). Evaluation of chloroplast genotypes of Korean cucumber cultivars (*Cucumis sativus* L.) using sdCAPS markers related to chilling tolerance. *Korean J. Hortic. Sci. Technol.* 31, 219–223. doi: 10.7235/hort.2013.12161

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi: 10.1038/35048692

Asghari, A., Mohammadi, S. A., Mohgaddam, M., and Mohammaddoost, H. R. (2007). Identification of QTLS controlling winter survival in *Brassica napus* using RAPD markers. *Biotechnol. Biotechnol. Equip.* 21, 413–416. doi: 10.1080/13102818.2007.10817485

Ashraf, M. (1994). Breeding for salinity tolerance in plants. *Crit. Rev. Plant Sci.* 13, 17–42. doi: 10.1080/713608051

Ashraf, M., Bokhari, M. H., and Mehmood, S. (1989). Effects of four different salts on germination and seedling growth of four *Brassica* species. *Biologia* 35, 173–187.

Ashraf, M., and McNeilly, T. (1990). Responses of four *Brassica* species to sodium chloride. *Environ. Exp. Bot.* 30, 475–487. doi: 10.1016/0098-8472(90)90013-T

Ashraf, M., McNeilly, T., and Nazir, M. (2001). Comparative salt tolerance of amphidiploid and diploid *Brassica* species. *Plant Sci.* 160, 683–689. doi: 10.1016/S0168-9452(00)00449-0

Babula, D., Kaczmarek, M., Barakat, A., Delseny, M., Quiros, C. F., and Sadowski, J. (2003). Chromosomal mapping of *Brassica oleracea* based on ESTs from *Arabidopsis thaliana*: complexity of the comparative map. *Mol. Genet. Genomics* 268, 656–665. doi: 10.1007/s00438-002-0782-2

Banerjee, S., Parasramka, M. A., and Sarkar, F. H. (2012). "Cellular, molecular and biological insight into chemopreventive and therapeutic potential of 3,3'-diindolylmethane (DIM)," in *Nutraceuticals and Cancer*, ed. F. H. Sarkar (New York, NY: Springer Netherlands), 111–133. doi: 10.1007/978-94-007-2630-7_6

Barret, P., Delourme, R., Foisset, N., and Renard, M. (1998). Development of a SCAR (sequence characterized amplified region) marker for molecular tagging

of the dwarf BREIZH (Bzh) gene in *Brassica napus* L. *Theor. Appl. Genet.* 97, 828–833. doi: 10.1007/s001220050962

Becker, H. C., Engqvist, G. M., and Karlssom, B. (1995). Comparison of rapeseed cultivars and resynthesized lines based on allozyme and RFLP markers. *Theor. Appl. Genet.* 91, 62–67. doi: 10.1007/BF00220859

Beecher, C. (1994). Cancer preventive properties of varieties of *Brassica Oleaceae*: a review. *Am. J. Clin. Nutr.* 59, 1166S–1170S.

Bell, C. J., and Ecker, J. R. (1994). Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* 19, 137–144. doi: 10.1006/geno.1994.1023

Bonilla, P., Dvorak, J., Mackill, D., Deal, K., and Gregorio, G. (2002). RFLP and SSLP mapping of salinity tolerance genes in chromosome 1 of rice (*Oryza sativa* L.) using recombinant inbred lines. *Philipp. Agric. Sci.* 85, 68–76.

Brown, G. G., Formanova, N., Jin, H., Wargachuk, R., Dendy, C., Patil, P., et al. (2003). The radish Rfo restorer gene of Ogura cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. *Plant J.* 35, 262–272. doi: 10.1046/j.1365-313X.2003.01799.x

Carvajal, M., Martinez, M., and Alcaraz, C. F. (1999). Physiological function of water channels as affected by salinity in roots of paprika pepper. *Physiol. Plant.* 105, 95–101. doi: 10.1034/j.1399-3054.1999.105115.x

Carvalhoa, I. S., Mirandab, I., and Pereira, H. (2006). Evaluation of oil composition of some crops suitable for human nutrition. *Ind. Crops Prod.* 24, 75–78. doi: 10.1016/j.indcrop.2006.03.005

Chalhoub, B., Budin, K., Aubourg, S., Belcram, H., Laffont, P., Delourme, R., et al. (2003). "Comparative genomic approaches for high-throughput physical and genetic mapping and gene cloning in oilseed rape (*Brassica napus*) and application to breeding," in *Proceedings of the 11th International Rapeseed Congress*, Vol. 1, Copenhagen, 78.

Chen, H., Cui, S., Fu, S., Gai, J., and Yu, D. (2008). Identification of quantitative trait loci associated with salt tolerance during seedling growth in soybean (*Glycine max* L.). *Aust. J. Agric. Res.* 59, 1086–1091. doi: 10.1071/AR08104

Delourme, R., Foisset, N., Horvais, R., Barret, P., Champagne, G., Cheung, W. Y., et al. (1998). Characterisation of the radish introgression carrying the Rfo restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 97, 129–134. doi: 10.1007/s001220050876

Diers, B. W., McVetty, P. B. E., and Osborn, T. C. (1996). Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (*Brassica napus* L.). *Crop Sci.* 36, 79–83. doi: 10.2135/cropsci1996.0011183X003600010014x

Diers, B. W., and Osborn, T. C. (1994). Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphisms. *Theor. Appl. Genet.* 88, 662–668. doi: 10.1007/BF01253968

Dion, Y., Gugel, R. K., Rakow, G. F. W., Seguin-Swartz, G., and Landry, B. S. (1995). RFLP mapping of resistance to the blackleg disease [causal agent, *Leptosphaeria maculans* (Desm.) Ces. et de Not. In canola (*Brassica napus* L). *Theor. Appl. Genet.* 91, 1190–1194. doi: 10.1007/BF00220928

Divisi, D., Tommaso, D. S., Salvemini, S., Garramone, M., and Crisci, R. (2006). Diet and cancer. *Acta. Biomed.* 77, 118–123.

Dreyer, F., Graichen, K., and Jung, C. (2001). A major quantitative trait locus for resistance to turnip yellows virus (TuYV, syn. beet western yellows virus, BWYV) in rapeseed. *Plant Breed.* 120, 457–462. doi: 10.1046/j.1439-0523.2001.00646.x

Dubcovsky, J., Maria, G. S., Epstein, E., Luo, M. C., and Dvorak, J. (1996). Mapping of the K+/Na+ discrimination locus Kna1 in wheat. *Theor. Appl. Genet.* 92, 448–454. doi: 10.1007/BF00223692

Ecke, W., Uzunova, M., and Weißleder, K. (1995). Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid systhesis and seed oil content. *Theor. Appl. Genet.* 91, 972–977. doi: 10.1007/BF00223908

Epstein, R. (1985). Extinction-induced resurgence: preliminary investigations and possible applications. *Psychol. Rec.* 35, 143–153.

Faircloth, B. C. (2008). msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Mol. Ecol. Resour.* 8, 92–94. doi: 10.1111/j.1471-8286.2007.01884.x

FAO. (2008). *FAO Land and Plant Nutrition Management Service*. Available at: http://www.fao.org/ag/agl/agll/spush

Ferreira, M. E., Rimmer, S. R., Williams, P. H., and Osborn, T. C. (1995a). Mapping loci controlling *Brassica napus* resistance to *Leptosphaeria maculans* under different screening conditions. *Phytopathology* 85, 213–217. doi: 10.1094/Phyto-85-213

Ferreira, M. E., Satagopan, J., Yandell, B. S., Williams, P. H., and Osborn, T. C. (1995b). Mapping loci controlling vernalization requirement and flowering time in *Brassica napus*. *Theor. Appl. Genet.* 90, 727–732. doi: 10.1007/BF00222140

Ferreira, M. E., Williams, P. H., and Osborn, T. C. (1995c). Mapping of a locus controlling resistance to *Albugo candida* in *Brassica napus* using molecular markers. *Phytopathology* 85, 218–220.

Finley, J. W., Sigrid-Keck, A., Robbins, R. J., and Hintze, K. J. (2005). Selenium enrichment of Broccoli: interactions between selenium and secondary plant compounds. *J. Nutr.* 135, 1236–1238.

Flowers, T. J., Koyama, M. L., Flowers, S. A., Sudhakar, C., Singh, K. P., and Yeo, A. R. (2000). QTL: their place in engineering tolerance of rice to salinity. *J. Exp. Bot.* 51, 99–106. doi: 10.1093/jexbot/51.342.99

Flowers, T. J., and Yeo, A. R. (1995). Breeding for salinity resistance in crop plants. Where next? *Aust. J. Plant Phys.* 22, 875–884. doi: 10.1071/PP9950875

Foisset, N., Delourme, R., Barret, P., and Renard, M. (1995). Molecular tagging of the dwarf Breizh (Bzh) gene in *Brassica napus*. *Theor. Appl. Genet.* 91, 756–761. doi: 10.1007/BF00220955

Fourmann, M., Barret, P., Renard, M., Pelletier, G., Delourme, R., and Brunel, D. (1998). The two genes homologous to *Arabidopsis* FAE1 co-segregate with the two loci governing erucic acid content in *Brassica napus*. *Theor. Appl. Genet.* 96, 852–858. doi: 10.1007/s001220050812

Francki, M., Snowdon, R. J., Wittkop, B., Rezaidad, A., Hasan, M., Lipsa, F., et al. (2010). Regional association analysis delineates a sequenced chromosome region influencing antinutritive seed meal compounds in oilseed rape. *Genome* 53, 917–928. doi: 10.1139/g10-052

Fray, M. J., Puangsomlee, P., Goodrich, J., Coupland, G., Evans, E. J., Arthur, A. E., et al. (1997). The genetics of stamenoid petal production in oilseed rape (*Brassica napus*) and equivalent variation in *Arabidopsis thaliana*. *Theor. Appl. Genet.* 94, 731–736. doi: 10.1007/s001220050472

Fu, F. Y., Liu, L. Z., Chai, Y. R., Chen, L., Yang, T., Jin, M. Y., et al. (2007). Localization of QTLs for seed color using recombinant inbred lines of *Brassica napus* in different environments. *Genome* 50, 840–854. doi: 10.1139/G07-068

Gorham, J., Britol, A., Young, E. M., and Wyn Jones, R. G. (1991). The presence of the enhanced K+/Na+ discrimination trait in diploid Triticum species. *Theor. Appl. Genet.* 82, 729–736. doi: 10.1007/BF00227318

Gorham, J., Hardy, C., Jones, R. G. W., Joppa, L. R., and Law, C. N. (1987). Chromosomal l location of a K/Na discrimination character in the D-genome of wheat. *Theor. Appl. Genet.* 74, 584–588. doi: 10.1007/BF00288856

Gorham, J., Jones, R. G. W., and Bristol, A. (1990). Partial characterization of the trait for enhanced K+-Na+ discrimination in the D-genome of wheat. *Planta* 180, 590–597. doi: 10.1007/BF02411458

Gregorio, G. B., Senadhira, D., Mendoza, R. D., Manigbas, N. L., Roxas, J. P., and Guerta, C. Q. (2002). Progress in breeding for salinity tolerance and associated abiotic stresses in rice. *Field Crops Res.* 76, 91–101. doi: 10.1016/S0378-4290(02)00031-X

Grieve, C. M., Suarez, D. L., and Shannon, M. C. (1999). Effect of saline irrigation water composition on selenium accumulation by wheat. *J. Plant Nutr.* 22, 1443–1450. doi: 10.1080/01904169909365725

Grist, S. A., Firgaira, F. A., and Morley, A. A. (1993). Dinucleotide repeat polymorphisms isolated by the polymerase chain reaction. *Biotechniques* 15, 304–309.

Gupta, V., Mukhopadhyay, A., Arumugam, N., Sodhi, Y. S., Pental, D., and Pradhan, A. K. (2004). Molecular tagging of erucic acid trait in oilseed mustard (*Brassica juncea*) by QTL mapping and single nucleotide polymorphisms in FAE1 gene. *Theor. Appl. Genet.* 108, 743–749. doi: 10.1007/s00122-003-1481-z

Gustafson, J. P., Badani, A. G., Snowdon, R. J., Wittkop, B., Lipsa, F. D., Baetzel, R., et al. (2006). Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (*Brassica napus*). *Genome* 49, 1499–1509. doi: 10.1139/g06-091

Halldén, C., Nilsson, N. O., Rading, I. M., and Säll, T. (1994). Evaluation of RFLP and RAPD markers in comparison of *Brassica napus* breeding lines. *Theor. Appl. Genet.* 88, 123–128. doi: 10.1007/BF00222404

Hamwieh, A., Tuyen, D., Cong, H., Benitez, E., Takahashi, R., and Xu, D. (2011). Identification and validation of a major QTL for salt tolerance in soybean. *Euphytica* 179, 451–459. doi: 10.1007/s10681-011-0347-8

Hasan, M., Friedt, W., Pons-Kühnemann, J., Freitag, N. M., Link, K., and Snowdon, R. J. (2008). Association of gene-linked SSR markers to seed glucosinolate content in oilseed rape (*Brassica napus* ssp. *napus*). *Theor. Appl. Genet.* 116, 1035–1049. doi: 10.1007/s00122-008-0733-3

Howell, P. M., Sharpe, A. G., and Lydiate, D. J. (2003). Homoeologous loci control the accumulation of seed glucosinolates in oilseed rape (*Brassica napus*). *Genome* 46, 454–460. doi: 10.1139/g03-028

Hu, J., Li, G., Struss, D., and Quiros, C. F. (1999). SCAR amd RAPD markers associated with 18-carbon fatty acid rapeseed, *Brassica napus*. *Plant Breed.* 118, 145–150. doi: 10.1046/j.1439-0523.1999.118002145.x

Hu, J., Quiros, C., Arus, P., Struss, D., and Robbelen, G. (1995). Mapping of a gene determining linolenic acid concentration in rapeseed with DNA-based markers. *Theor. Appl. Genet.* 90, 258–262. doi: 10.1007/BF00222211

Hu, X., Sullivan-Gilbert, M., Gupta, M., and Thompson, S. A. (2006). Mapping of the loci controlling oleic and linolenic acid contents and development of fad2 and fad3 allele-specific markers in canola (*Brassica napus* L.). *Theor. Appl. Genet.* 113, 497–507. doi: 10.1007/s00122-006-0315-1

Huang, S. B., Spielmeyer, W., Lagudah, E. S., James, R. A., Platten, J. D., Dennis, E. S., et al. (2006). A sodium transporter (HKT7) is a candidate for Nax1, a gene for salt tolerance in durum wheat. *Plant Physiol.* 142, 1718–1727.

Huang, S., Spielmeyer, W., Lagudah, E. S., and Munns, R. (2008). Comparative mapping of HKT genes in wheat, barley, and rice, key determinants of Na+ transport, and salt tolerance. *J. Exp. Bot.* 59, 927–937. doi: 10.1093/jxb/ern033

Im, C. H., Kim, M. K., Kim, K. H., Cho, S. J., Lee, J. J., Joung, W. K., et al. (2014a). Breeding of *Pleurotus eryngii* with a high temperature tolerance trait. *J. Mushrooms* 12, 187–192. doi: 10.4489/KJM.2013.41.2.91

Im, C. H., Kim, K. H., Je, H. J., Ali, A., Kim, M. K., Joung, W. K., et al. (2014b). Multiplex simple sequence repeat (SSR) markers discriminating *Pleurotus eryngii* cultivar. *Korean J. Mycol.* 42, 159–164. doi: 10.4489/KJM.2014.42.2.159

Imai, R., Koizuka, N., Fujimoto, H., Hayakawa, T., Sakai, T., and Imamura, J. (2003). Delimitation of the fertility restorer locus Rfk1 to a 43-kb contig in Kosena radish (*Raphanus sativus* L.). *Mol. Genet. Genomics* 269, 388–394. doi: 10.1007/s00438-003-0847-x

Islam, M. R., Salam, M. A., Hassan, L., Collard, B. C. Y., Singh, R. K., and Gregorio, G. B. (2011). QTL mapping for salinity tolerance in at seedling stage in rice. *Emirates J. Food Agric.* 23, 137–146. doi: 10.9755/ejfa.v23i2.6348

James, R. A., Davenport, R. J., and Munns, R. (2006). Physiological characterization of two genes for Na+ exclusion in durum wheat, Nax1 and Nax2. *Plant Physiol.* 142, 1537–1547. doi: 10.1104/pp.106.086538

Janeja, H. S., Banga, S. S., and Lakshmikumaran, M. (2003). Identification of AFLP markers linked to fertility restorer genes for tournefortii cytoplasmic male-sterility system in *Brassica napus*. *Theor. Appl. Genet.* 107, 148–154. doi: 10.1007/s00122-003-1225-0

Jean, M., Brown, G. G., and Landry, B. S. (1998). Targeted mapping approaches to identify DNA markers linked to the Rfp1 restorer gene for the 'Polima' CMS of canola (*Brassica napus* L.). *Theor. Appl. Genet.* 97, 431–438. doi: 10.1007/s001220050913

Kagale, S., Divi, U. K., Krochko, J. E., Keller, W. A., and Krishna, P. (2007). Brassinosteroid confers tolerance in *Arabidopsis thaliana* and *Brassica napus* to a range of abiotic stresses. *Planta* 225, 353–364. doi: 10.1007/s00425-006-0361-6

Kole, C., Quijada, P., Michael, S. D., Amasino, R., and Osborn, T. C. (2001). Evidence for homology of flowering-time genes VFR2 from *Brassica rapa* and FLC from *Arabidopsis thaliana*. *Theor. Appl. Genet.* 102, 425–430. doi: 10.1007/s001220051663

Kole, C., Thormann, C. E., Karlsson, B. H., Palta, J. P., Gaffney, P., Yandell, B., et al. (2002a). Comparative mapping of loci controlling winter survival and related traits in oilseed *Brassica rapa* and *B. napus*. *Mol. Breed.* 9, 201–210. doi: 10.1023/A:1019759512347

Kole, C., Williams, P. H., Rimmer, S. R., and Osborn, T. C. (2002b). Linkage mapping of genes controlling resistance to white rust (*Albugo candida*) in *Brassica rapa* (syn. campestris) and comparative mapping to *Brassica napus* and *Arabidopsis thaliana*. *Genome* 45, 22–27. doi: 10.1139/g01-123

Kowalski, S. P., Lan, T. H., Feldmann, K. A., and Paterson, A. H. (1994). Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* 138, 499–510.

Kumar, K., Kumar, M., Kim, S. R., Ryu, H., and Cho, Y. G. (2013). Insights into genomics of salt stress response in rice. *Rice (N. Y.)* 6, 27. doi: 10.1186/1939-8433-6-27

Kumar, M. (2013). Crop plants and abiotic stresses. *J. Biomol. Res. Ther.* 3, e125. doi: 10.4172/2167-7956.1000e125

Kumar, M. (2015). Detection of genetic variation in crop plants. *J. Biomol. Res. Ther.* 4, e140. doi: 10.4172/2167-7956.1000e140

Kumar, M., Kim, S. R., Sharma, P. C., and Pareek, A. (2015). Simple and efficient way to detect small polymorphic bands in plants. *Genomics Data* 5, 218–222. doi: 10.1016/j.gdata.2015.06.006

Kumar, M., Lee, S. C., Kim, J. Y., Kim, S. J., Aye, S. S., and Kim, S. R. (2014). Over-expression of dehydrin gene, OsDhn1, improves drought and salt stress tolerance through scavenging of reactive oxygen species in rice (*Oryza sativa* L.). *J. Plant Biol.* 57, 383–393. doi: 10.1007/s12374-014-0487-1

Lagercrantz, U. (1998). Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that Brassica genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150, 1217–1228.

Lagercrantz, U., Ellegren, H., and Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Res.* 21, 1111–1115. doi: 10.1093/nar/21.5.1111

Lagercrantz, U., Putterill, J., Coupland, G., and Lydiate, D. (1996). Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. *Plant J.* 9, 13–20. doi: 10.1046/j.1365-313X.1996.09010013.x

Lan, T. H., and Paterson, A. H. (2001). Comparative mapping of QTLs determining the plant size of *Brassica oleracea*. *Theor. Appl. Genet.* 103, 383–397. doi: 10.1007/s001220100615

Landry, B. S., Hubert, N., Etoh, T., Harada, J. J., and Lincoln, S. E. (1991). A genetic map for *Brassica napus* based on restriction fragment length polymorphisms detected with expressed DNA sequences. *Genome* 34, 543–552. doi: 10.1139/g91-084

Lang, N. T., Yanagihara, S. S., and Buu, B. C. (2001). A microsatellite marker for a gene conferring salt tolerance on rice at the vegetative and reproductive stages. *Sabrao J. Breed. Genet.* 33, 1–10.

Lee, G. P., Baek, N. K., and Park, K. W. (2002). Genetic mapping of resistant genes in *Brassica pekinensis* against *Plasmodiophora brassicae* race 6. *Plant Pathol. J.* 18, 266–270. doi: 10.5423/PPJ.2002.18.5.266

Lee, G. J., Carter, T. E. Jr., Villagarcia, M. R., Li, Z., Zhou, X., Gibbs, M. O., et al. (2004). A major QTL conditioning salt tolerance in S-100 soybean and descendent cultivars. *Theor. Appl. Genet.* 109, 1610–1619. doi: 10.1007/s00122-004-1783-9

Lexer, C., Welch, M. E., Durphy, J. L., and Rieseberg, L. H. (2003). Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: implications for the origin of *Helianthus paradoxus*, a diploid hybrid species. *Mol. Ecol.* 12, 1225–1235. doi: 10.1046/j.1365-294X.2003.01803.x

Lin, H. X., Zhu, M. Z., Yano, M., Gao, J. P., Liang, Z. W., Su, W. A., et al. (2004). QTLs for Na+ and K+ uptake of the shoots and roots controlling rice salt tolerance. *Theor. Appl. Genet.* 108, 253–260. doi: 10.1007/s00122-003-1421-y

Lindsay, M. P., Lagudah, E. S., Hare, R. A., and Munns, R. (2004). A locus for sodium exclusion (*Nax1*), a trait for salt tolerance, mapped in durum wheat. *Funct. Plant Biol.* 31, 1105–1114. doi: 10.1071/FP04111

Litt, M., and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Amr. J. Hum. Genet.* 44, 397–401.

Lu, G., Cao, J., Yu, X., Xiang, X., and Chen, H. (2008). Mapping QTLs for root morphological traits in *Brassica rapa* L. based on AFLP and RAPD markers. *J. Appl. Genet.* 49, 23–31. doi: 10.1007/BF03195245

Luo, M. C., Dubcovsky, J., Goyal, S., and Dvorak, J. (1996). Engineering of interstitial foreign chromosome segments containing the K+/Na+ selectivity gene Kna1 by sequential homoeologous recombination in durum wheat. *Theor. Appl. Genet.* 93, 1180–1184. doi: 10.1007/BF00230144

Martins, W. S., Lucas, D. C. S., Neves, K. F. S., and Bertioli, D. J. (2009). WebSat - A Web Software for MicroSatellite marker development. *Bioinformation* 3, 282–283. doi: 10.6026/97320630003282

Morgante, M., Hanafey, H., and Powell, W. (2002). Microsatellites are preferentially associated with nonrepetitive DNA in plant genome. *Nature Genet.* 30, 194–200. doi: 10.1038/ng822

Mudunuri, S. B., and Nagarajaram, H. A. (2007). IMEx: imperfect microsatellite extractor. *Bioinformatics* 23, 1181–1187. doi: 10.1093/bioinformatics/btm097

Mukhlesur, R., Mcvetty, P. B. E., and Li, G. (2007). Development of SRAP, SNP and Multiplexed SCAR molecular markers for the major seed coat color gene in *Brassica rapa* L. *Theor. Appl. Genet.* 115, 1101–1107. doi: 10.1007/s00122-007-0636-8

Mullis, K., and Faloona, F. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 44, 224–249.

Nagaharu, U. (1935). Genome analysis in Brassica with special reference to the experimental formation of *Brassica napus* and peculiar mode of fertilization. *Japan J. Bot.* 7, 389–452.

Okazaki, K., Sakamoto, K., Kikuchi, R., Saito, A., Togashi, E., Kuginuki, Y., et al. (2007). Mapping and characterization of FLC homologs and QTL analysis of flowering time in *Brassica oleracea*. *Theor. Appl. Genet.* 114, 595–608. doi: 10.1007/s00122-006-0460-6

O'Neill, R., Snowdon, R. J., and Köhler, W. (2003). Population genetics aspects of biodiversity. *Prog. Bot.* 64, 115–137. doi: 10.1007/978-3-642-55819-1_7

Padmaja, K. L., Arumugam, N., Gupta, V., Mukhopadhyay, A., Sodhi, Y. S., Pental, D., et al. (2005). Mapping and tagging of seed coat colour and the identifcation of markers for marker-assisted manipulation of the trait in *Brassica juncea* microsatellite. *Theor. Appl. Genet.* 111, 8–14. doi: 10.1007/s00122-005-1933-8

Panjabi, P., Jagannath, A., Bisht, N. C., Padmaja, K. L., Sharma, S., Gupta, V., et al. (2008). Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using Intron Polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C *Brassica* genomes. *BMC Genomics* 9:113. doi: 10.1186/1471-2164-9-113

Parkin, I. A. P., Gulden, S. M., Sharpe, A. G., Lukens, L., Trick, M., Osborn, T. C., et al. (2005). Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics* 171, 765–781. doi: 10.1534/genetics.105.042093

Pflieger, S., Lefebvre, V., and Causse, M. (2001). The candidate gene approach in plant genetics: a review. *Mol. Breed.* 7, 275–291. doi: 10.1023/A:1011605013259

Pilet, M. L., Delourme, R., Foisset, N., and Renard, M. (1998a). Identification of loci contributing to quantitative resistance to blackleg disease, causal agent *Leptsosphaeria maculans* (Desm.) Ces. et de Not., in winter rapseed (*Brassica napus* L.). *Theor. Appl. Genet.* 96, 23–30. doi: 10.1007/s001220050704

Pilet, M. L., Delourme, R., Foisset, N., and Renard, M. (1998b). Identification of QTL involved in field resistance to light leaf spot (*Pyrenopeziza brassicae*) and blackleg resistance (*Leptosphaeria maculans*) in winter rapeseed (*Brassica napus* L.). *Theor. Appl. Genet.* 97, 398–406. doi: 10.1007/s001220050909

Pilet, M. L., Duplan, G., Archipiano, H., Barret, P., Baron, C., Horvais, R., et al. (2001). Stability of QTL for field resistance to blackleg across two genetic backgrounds in oilseed rape. *Crop Sci.* 41, 197–205. doi: 10.2135/cropsci2001.411197x

Plieske, J., and Struss, D. (2001a). Microsatellite markers for genome analysis in *Brassica*. I. development in *Brassica napus* and abundance in *Brassicaceae species*. *Theor. Appl. Genet.* 102, 689–694. doi: 10.1007/s001220051698

Plieske, J., and Struss, D. (2001b). STS markers linked to Phoma resistance genes of the *Brassica* B-genome revealed sequence homology between *Brassica nigra* and *Brassica napus*. *Theor. Appl. Genet.* 102, 483–488. doi: 10.1007/s001220051671

Putterill, J., Robson, F., Lee, K., and Coupland, G. (1995). The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Qu, C., Hasan, M., Lu, K., Liu, L., Zhang, K., Fu, F., et al. (2015). Identification of QTL for seed coat colour and oil content in *Brassica napus* by association mapping using SSR markers. *Can. J. Plant Sci.* 95, 387–395. doi: 10.4141/cjps2013-411

Ramchiary, N., Padmaja, K. L., Sharma, S., Gupta, V., Sodhi, Y. S., Mukhopadhyay, A., et al. (2007). Mapping of yield influencing QTL in *Brassica juncea*: implications for breeding of a major oilseed crop of dryland areas. *Theor. Appl. Genet.* 115, 807–817. doi: 10.1007/s00122-007-0610-5

Ren, Z. H., Gao, J. P., Li, L. G., Cai, X. L., Huang, W., Chao, D. Y., et al. (2005). A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nat. Genet.* 37, 1141–1146. doi: 10.1038/ng1643

Rezaeizad, A., Wittkop, B., Snowdon, R., Hasan, M., Mohammadi, V., Zali, A., et al. (2011). Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. *Euphytica* 177, 335–342. doi: 10.1007/s10681-010-0231-y

Riaz, A., Li, G., Quresh, Z., Swati, M. S., and Quiros, C. F. (2001). Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. *Plant Breed.* 120, 411–415. doi: 10.1046/j.1439-0523.2001.00636.x

Robert, L. S., Robson, F., Sharpe, A., Lydiate, D., and Coupland, G. (1998). Conserved structure and function of the *Arabidopsis* flowering time gene CONSTANS in *Brassica napus*. *Plant Mol. Biol.* 37, 763–772. doi: 10.1023/A:1006064514311

Sabouri, H. (2009). QTL detection of rice grain quality traits by microsatellite markers using an indica rice (*Oryza sativa* L.) combination. *J. Genet.* 88, 81–85. doi: 10.1007/s12041-009-0011-4

Sabouri, H., and Sabouri, A. (2008). New evidence of QTLs attributed to salinity tolerance in rice. *Afr. J. Biotechnol.* 7, 4376–4383. doi: 10.1007/s11033-013-2561-z

Seyis, F., Snowdon, R. J., Lühs, W., and Friedt, W. (2003). Molecular characterization of novel resynthesized rapeseed (*Brassica napus*) lines and analysis of their genetic diversity in comparison with spring rapeseed cultivars. *Plant Breed.* 122, 473–478. doi: 10.1111/j.1439-0523.2003.00859.x

Sharma, R., Mishra, M., Gupta, B., Parsania, C., Singla-Pareek, S. L., and Pareek, A. (2015). De novo assembly and characterization of stress transcriptome in a salinity tolerant variety CS52 of *Brassica juncea*. *PLoS ONE* 10:e0126783. doi: 10.1371/journal.pone.0126783

Shavrukov, Y., Gupta, N. K., Miyazaki, J., Baho, M. N., Chalmers, K. J., Tester, M., et al. (2010). HvNax3-a locus controlling shoot sodium exclusion derived from wild barley (*Hordeum vulgare* ssp. *spontaneum*). *Funct. Integr. Genomics* 10, 277–291. doi: 10.1007/s10142-009-0153-8

Shi, J., Huang, S., Zhan, J., Yu, J., Wang, X., Hua, W., et al. (2014). Genome-wide microsatellite characterization and marker development in the sequenced *Brassica* crop species. *DNA Res.* 21, 53–68. doi: 10.1093/dnares/dst040

Singan, V., and Colbourne, J. K. (2005). *MicrosatDesign is a Pipeline for Transforming Sequencer Trace Files into DNA Markers*. CGB Technical Report 2005-01. Bloomington, IN: The Center for Genomics and Bioinformatics, Indiana University.

Singh, R. K., Mishra, B., and Vandna, J. (2001). Segregations for alkalinity tolerance in three rice crosses. *SABRAO J.* 33, 31–34.

Slocum, M. K., Figdore, S. S., Kennard, W. C., Suzuki, J. Y., and Osborn, T. C. (1990). Linkage arrangement of restriction fragment length polymorphisms in *Brassica oleracea*. *Theor. Appl. Genet.* 80, 57–64. doi: 10.1007/BF00224016

Smith, D. N., and Devey, M. E. (1994). Occurrence and inheritance of microsatellites in *Pinus radia*. *Genome* 37, 977–983. doi: 10.1139/g94-138

Smykal, P., Weising, K., Nybom, H., Wolf, K., and Kahl, G. (2006). DNA fingerprinting in plants. principles, methods and applications, 2nd edition. *Biol. Plant.* 50, 799. doi: 10.1007/s10535-006-0138-6

Somers, D. J., Friesen, K. R. D., and Rakow, G. (1998). Identifcation of molecular markers associated with linoleic acid desaturation in *Brassica napus*. *Theor. Appl. Genet.* 96, 897–903. doi: 10.1007/s001220050817

Somers, D. J., Rakow, G., Prabhu, V. K., and Friesen, K. R. D. (2001). Identification of a major gene and RAPD markers for yellow seed coat color in *B. napus*. *Genome* 44, 1077–1082. doi: 10.1139/gen-44-6-1077

Song, K. M., Suzuki, J. Y., Slocum, M. K., Williams, P. H., and Osborn, T. C. (1991). A linkage map of *Brassica rapa* (syn. campestris) based on restriction fragment length polymorphisms. *Theor. Appl. Genet.* 82, 296–304. doi: 10.1007/BF02190615

Sotelo, T., Soengas, P., Velasco, P., Rodríguez, V. M., and Cartea, M. E. (2014). Identification of metabolic qtls and candidate genes for glucosinolate synthesis in *Brassica oleracea* leaves, seeds and flower buds. *PLoS ONE* 9:e91428. doi: 10.1371/journal.pone.0091428

Suwabe, K., Iketani, H., Nunome, T., Kage, T., and Hirai, M. (2002). Isolation and characterization of microsatellites in *Brassica rapa* L. *Theor. Appl. Genet.* 104, 1092–1098. doi: 10.1007/s00122-002-0875-7

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Fujimura, M., Nunome, T., et al. (2003). Identification of two loci for resistance to clubroot

(*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. *Theor. Appl. Genet.* 107, 997–1002. doi: 10.1007/s00122-003-1309-x

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fujimura, M., et al. (2006). Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. *Genetics* 173, 309–319. doi: 10.1534/genetics.104.038968

Tanhuanpää, P., and Schulman, A. (2002). Mapping of genes affecting linolenic acid content in *Brassica rapa* ssp. oleifera. *Mol. Breed.* 10, 51–62. doi: 10.1023/A:1020357211089

Thiel, T., Michalek, W., Varshney, R. K., and Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106, 411–422. doi: 10.1007/s00122-002-1031-0

Thomson, M. J., De, O. M., Egdane, J., Rahman, M. A., Sajise, A. G., Adorada, D. L., et al. (2010). Characterizing the *Saltol* quantitative trait locus for salinity tolerance in rice. *Rice* 3, 148–160. doi: 10.1007/s12284-010-9053-8

Thormann, C. E., Ferreira, M. E., Camargo, L. E. A., Tivang, J. G., and Osborn, T. C. (1994). Comparison of RFLP and RAPD markers to estimating genetic relationships within and among cruciferous species. *Theor. Appl. Genet.* 88, 973–980. doi: 10.1007/BF00220804

Thurston, M., and Field, D. (2005). *Msatminer: Detection and Characterization of Microsatellites*. Oxford: Centre for Ecology & Hydrology.

Uzunova, M. I., and Ecke, W. (1999). Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). *Plant Breeding* 118, 323–326. doi: 10.1046/j.1439-0523.1999.00371.x

Uzunova, M., Ecke, W., Weissleder, K., and Robbelen, G. (1995). Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. *Theor. Appl. Genet.* 90, 194–204. doi: 10.1007/BF00222202

Walsh, J. A., Sharpe, A. G., Jenner, C. E., and Lydiate, D. J. (1999). Characterisation of resistance to turnip mosaic virus in oilseed rape (*Brassica napus*) and genetic mapping of TuRB01. *Theor. Appl. Genet.* 99, 1149–1154. doi: 10.1007/s001220051319

Wang, J., Drayton, M. C., George, J., Cogan, N. O., Baillie, R. C., Hand, M. L., et al. (2010). Identification of genetic factors influencing salt stress tolerance in white clover (*Trifolium repens* L.) by QTL analysis. *Theor. Appl. Genet.* 120, 607–619. doi: 10.1007/s00122-009-1179-y

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S., et al. (2011). The genome of the mesopolyploid crop species *Brassica rapa*. *Nat. Genet.* 43, 1035–1039. doi: 10.1038/ng.919

Weising, K., Nybom, H., Wolff, K., and Kahl, G. (2005). *DNA Fingerprinting in Plants: Principles, Methods, and Applications,* 2nd Edn. Boca Raton, FL: CRC Press Taylor & Francis Group.

Weiss, E. (1983). *Oilseed Crops, Tropical Agriculture Series*. London: Longman Scientific and Technical, 530–564.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* 18, 6531–6535. doi: 10.1093/nar/18.22.6531

Xiao, S., Xu, J., Li, Y., Zhang, L., Shi, S., Shi, S., et al. (2007). Generation and mapping of SCAR and CAPS markers linked to the seed coat color gene in *Brassica napus* using a genome-walking technique. *Genome* 50, 611–618. doi: 10.1139/G07-044

Xue, D. W., Zhou, M. X., Zhang, X. Q., Chen, S., Wei, K., Zeng, F. R., et al. (2010). Identification of QTLs for yield and yield components of barley under different growth conditions. *J. Zhejiang Univ. Sci. B* 11, 169–176. doi: 10.1631/jzus.B0900332

Yamaguchi, T., and Blumwald, E. (2005). Developing salt-tolerant crop plants: challenges and opportunities. *Trends Plant Sci.* 10, 615–620. doi: 10.1016/j.tplants.2005.10.002

Yao, M. Z., Wang, J. F., Chen, H. Y., Zhai, H. Q., and Zhang, H. S. (2005). Inheritance and QTL mapping of salt tolerance in rice. *Rice Sci.* 12, 25–32.

Yeo, A. R. (1998). Molecular biology of salt tolerance in the context of whole-plant physiology. *J. Exp. Bot.* 49, 915–929. doi: 10.1093/jxb/49.323.915

Yeo, A. R., and Flowers, T. J. (1986). "Salinity resistance in rice and a pyramiding approach to breeding varieties for saline soils," in *Plant Growth, Drought, and Salinity*, eds N. C. Turner and J. B. Passioura (Melbourne, VIC: CSIRO), 161–173.

Yu, H., and Li, Q. (2008). Exploiting EST databases for the development and characterization of EST-SSRs in the Pacific Oyster (*Crassostrea gigas*). *J. Hered.* 99, 208–214. doi: 10.1093/jhered/esm124

Yu, J. K., La Rota, M., Kantety, R. V., and Sorrells, M. E. (2004). EST derived SSR markers for comparative mapping in wheat and rice. *Mol. Genet. Genomics* 271, 742–751. doi: 10.1007/s00438-004-1027-3

Yu, J., Zhao, M., Wang, X., Tong, C., Huang, S., Tehrim, S., et al. (2013). Bolbase: a comprehensive genomics database for *Brassica oleracea*. *BMC Genomics* 14:664. doi: 10.1186/1471-2164-14-664

Zhang, Y., Sledge, M. K., and Bouton, J. H. (2007). Genome mapping of white clover (*Trifolium repens* L.) and comparative analysis within the Trifolieae using cross-species SSR markers. *Theor. Appl. Genet.* 114, 1367–1378. doi: 10.1007/s00122-007-0523-3

Zhao, J., and Meng, J. (2003). Genetic analysis of loci associated with partial resistance to *Sclerotinia sclerotiorum* in rapeseed (*Brassica napus L.*). *Theor. Appl. Genet.* 106, 759–764. doi: 10.1007/s00122-002-1171-2

Zhou, M., Johnson, P., Zhou, G., Li, C., and Lance, R. C. M. (2012). Quantitative trait loci for waterlogging tolerance in a barley cross of franklin x YuYaoXiangTian Erleng and the relationship between waterlogging and salinity tolerance. *Crop Sci.* 52, 2082–2088. doi: 10.2135/cropsci2012.01.0008

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kumar, Choi, Kumari, Pareek and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 February 2016
doi: 10.3389/fpls.2016.00220

Arabidopsis AtDjA3 Null Mutant Shows Increased Sensitivity to Abscisic Acid, Salt, and Osmotic Stress in Germination and Post-germination Stages

*Silvia Salas-Muñoz, Aída A. Rodríguez-Hernández, Maria A. Ortega-Amaro, Fatima B. Salazar-Badillo and Juan F. Jiménez-Bremont**

Laboratorio de Biotecnología Molecular de Plantas, División de Biología Molecular, Instituto Potosino de Investigación Científica y Tecnológica, San Luis Potosí, México

OPEN ACCESS

Edited by:
Susana Araújo, Instituto de Tecnologia Química e Biológica – Universidade Nova de Lisboa, Portugal

Reviewed by:
Anca Macovei, International Rice Research Institute, Philippines
Ji-Hong Liu, Huazhong Agricultural University, China

***Correspondence:**
Juan F. Jiménez-Bremont
jbremont@ipicyt.edu.mx

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *08 September 2015*
Accepted: *09 February 2016*
Published: *25 February 2016*

Citation:
Salas-Muñoz S, Rodríguez-Hernández AA, Ortega-Amaro MA, Salazar-Badillo FB and Jiménez-Bremont JF (2016) Arabidopsis AtDjA3 Null Mutant Shows Increased Sensitivity to Abscisic Acid, Salt, and Osmotic Stress in Germination and Post-germination Stages. Front. Plant Sci. 7:220. doi: 10.3389/fpls.2016.00220

DnaJ proteins are essential co-chaperones involved in abiotic and biotic stress responses. *Arabidopsis AtDjA3* gene encodes a molecular co-chaperone of 420 amino acids, which belongs to the J-protein family. In this study, we report the functional characterization of the *AtDjA3* gene using the *Arabidopsis* knockout line designated *j3* and the *35S::AtDjA3* overexpression lines. Loss of *AtDjA3* function was associated with small seed production. In fact, *j3* mutant seeds showed a reduction of 24% in seed weight compared to Col-0 seeds. Expression analysis showed that the *AtDjA3* gene was modulated in response to NaCl, glucose, and abscisic acid (ABA). The *j3* line had increased sensitivity to NaCl and glucose treatments in the germination and cotyledon development in comparison to parental Col-0. Furthermore, the *j3* mutant line exhibited higher ABA sensitivity in comparison to parental Col-0 and *35S::AtDjA3* overexpression lines. In addition, we examined the expression of *ABI3* gene, which is a central regulator in ABA signaling, in *j3* mutant and *35S::AtDjA3* overexpression lines. Under 5 μM ABA treatment at 24 h, *j3* mutant seedlings displayed higher *ABI3* expression, whereas in *35S::AtDjA3* overexpression lines, *ABI3* gene expression was repressed. Taken together, these results demonstrate that the *AtDjA3* gene is involved in seed development and abiotic stress tolerance.

Keywords: ***Arabidopsis thaliana*, *AtDjA3*, abscisic acid, abiotic stress, heat shock proteins, J-protein**

INTRODUCTION

Seed germination and seedling establishment are the most critical stages of survival during the life cycle of an individual plant (Daszkowska-Golec, 2011). Seeds are exposed to a wide range of unfavorable environmental conditions that induce stress, and therefore have a negative impact on germination, growth, and development (Rao et al., 2006). As a result, seeds have developed defense mechanisms that allow them to tolerate and respond rapidly to unfavorable conditions (Koornneef et al., 2002; Vallejo et al., 2010).

Heat Shock Proteins (HSP) are accumulated during abiotic stress as a defense mechanism, and at the later stages of seed development appear to play a protective role in desiccation tolerance

(Wehmeyer and Vierling, 2000; Koornneef et al., 2002). HSPs are involved in a variety of cellular processes including protein folding, assembly of oligomeric proteins, transport of proteins across membranes, stabilization of polypeptide strands and membranes, and prevention of protein inactivation (Vierling, 1991; Wang et al., 2004; Xue et al., 2010). In plants, HSPs are classified into five classes according to their molecular weight: HSP100, HSP90, HSP70, HSP60, and small HSP (sHSP; Wang et al., 2004).

HSP40 proteins, also referred to as DnaJ or J-proteins, are co-chaperones of the HSP70 machine. J-proteins are key players in stimulating HSP70 ATPase activity, thereby stabilizing its interaction with client proteins (Bukau and Horwich, 1998; Walsh et al., 2004). The J-proteins contain in the N-terminus region a highly conserved domain of approximately 70 amino acids, known as the J domain. This domain consists of four α-helices comprising two short helices (I and IV) and two tightly packed anti-parallel helices (II and III) linked by a loop region that contains a highly conserved tri-peptide (HPD: histidine-proline-aspartic acid), which is required for interaction with HSP70 proteins (Wall et al., 1994; Rajan and D'Silva, 2009). Adjacent to J domain is a characteristic glycine and phenylalanine (G/F) rich region. It has been proposed that this region serves as a flexible linker region and controls the specificity of J-protein functions (Craig et al., 2006). After the G/F-rich region there is a cysteine-rich region which forms a type I zinc-finger domain, which contains four repeated motifs (CXXCXGXG). This domain is essential for binding to unfolded protein and assists HSP70 with protein folding (Bukau and Horwich, 1998; Lu and Cyr, 1998). Finally, the C-terminal region, which is less conserved, is important for providing specificity for HSP70 J-protein machinery (Shi et al., 2005). Classification of J-proteins include the type I or A proteins, which present all the characteristic domains or regions; type II or B proteins, that lack the zinc-finger domain, and type III or C proteins that only contain domain J (Rajan and D'Silva, 2009). In the *Arabidopsis thaliana* genome, 116 J-proteins and four J-like proteins have been identified; of which eight belong to type I, 16 to type II, and 92 to type III (Rajan and D'Silva, 2009). The *A. thaliana* AtDjA3 protein belongs to the type I classification. In plants, J-proteins are induced under different stress conditions. *AtDjA3* gene is expressed in roots, stems, leaves, flower buds, flowers, and siliques, and its expression can be induced by heat, cold, and drought stress (Li et al., 2005, 2007), and also under saline conditions with alkaline pH (Yang et al., 2010).

In the present study, we deepen our understanding of *AtDjA3* gene under salt and osmotic stress, and the application of phytohormone abscisic acid (ABA). Expression analysis of the *AtDjA3* gene in *Arabidopsis* seedlings revealed that its expression is modulated by NaCl, glucose, and ABA. For the molecular characterization of *AtDjA3* gene, we analyzed the *Atdja3*-null (*j3*) mutant and *35S::AtDjA3* overexpression lines. Our results reveal that *j3* loss-of-function mutant produces small seeds that are less tolerance to salt and osmotic stress, reflected by a reduced germination rate and lower percentage of green cotyledons in comparison to the Col-0 and *35S::AtDjA3* overexpression lines. In addition, the *j3* mutant line shows more sensitivity to exogenous ABA. Our results suggest that *AtDjA3* gene plays a role in abiotic stress tolerance.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The *Atdja3* mutant line (*j3*) and transgenic lines (*OvJ3*-8 and -14) used in this study were generated in the *Arabidopsis thaliana* ecotype Columbia 0 (Col-0) background. The T-DNA insertion line (Salk_132923) for the *AtDjA3* gene (At3g44110) was obtained from the Salk Institute Genomic Analysis Laboratory[1] (Alonso et al., 2003). The seeds used for all experiments were harvested at the same time. The seeds of *A. thaliana* Col-0, *Atdja3*-null mutant line (*j3*) and *35S::AtDjA3* overexpression lines (*OvJ3*) were sterilized with 20% (v/v) commercial sodium hypochlorite (6% free chlorine) solution for 5 min, and rinsed five times in sterile distilled water. Aseptic seeds were germinated and grown on agar plates containing Murashige and Skoog (MS) 0.5x medium supplemented with 7 g/L phytagel, and 1.5% sucrose (Murashige and Skoog, 1962). Plates were kept at 4°C for 3 days, and then incubated at 22 ± 2°C for 10 days in a growth chamber under a 16 h light (13,000 luxes)/8 h dark photoperiod. Plants were grown to maturity in soil pots, with a mixture of Sunshine Mix #3 commercial substrate, perlite and vermiculite (3:1:1), in a growth chamber at 22 ± 2°C with a 16 h light (13,000 luxes)/8 h dark photoperiod.

Nucleic Acids Isolation and cDNA Synthesis

Genomic DNA was isolated from *A. thaliana* WT (Col-0) and T-DNA insertion mutant line plants using the method described by Murray and Thompson (1980). RNA extraction from Col-0, *Atdja3*-null mutant line and *35S::AtDjA3* overexpression lines was performed using the Concert Plant RNA Reagent (Invitrogen, Carlsbad, CA, USA) by following the manufacturer's instructions; samples were stored at −70°C until analysis. For the removal of contaminating genomic DNA, RNA samples were treated with DNase I (Invitrogen, Carlsbad, CA, USA). Synthesis of cDNA was carried out with the Super Script II Reverse Transcriptase enzyme (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. The cDNAs were stored at −20°C for subsequent use.

Identification of the T-DNA Insertional Mutant Line

Seeds of T-DNA insertion mutant line (Salk_132923) were germinated in plates containing MS 0.5x medium; after 10 days they were transferred to soil pots. For genotype analysis PCR assays were performed on genomic DNA isolated from 3-week-old plants using the T-DNA left border (LB) oligonucleotide and gene-specific oligonucleotides designed flanking the T-DNA. The oligonucleotides used were:

[1]http://signal.salk.edu

FwSalk_132923 5′-CTTGAAGGTATCTCTTGAGGATGTGT ACC-3′, RvSalk_132923 5′-GACGATGCATCTGAATACGTACC AGG-3′, and FwLB 5′- AGCAAGCGGTCCACGCTGGTTT-3′.

Generation of *A. thaliana 35S::AtDjA3* Overexpression Lines

The *AtDjA3* open reading frame (At3g44110 GenBank ID: 823531) was amplified from a cDNA sample of *Arabidopsis* seedlings using the Hot Start High-Fidelity Polymerase Kit (Qiagen, USA), and the following oligonucleotides: FwAtDjA3 5′-GGCGAAAAGATGTTCGGTAGAGG-3′ and RvAtDjA3 5′-GTCTCTCTAAGGAGTTACTTACTGC-3′. The amplified product of 1,263 bp was cloned into the pCR8/GW/TOPO vector (Invitrogen, Carlsbad, CA, USA). The cloned products were sequenced using the M13 oligonucleotide in an ABIPRISM 377 DNA automated sequencer (Perkin Elmer, USA). The sequenced entry clone was recombined into the pMDC32 destination vector (Curtis and Grossniklaus, 2003), by site-specific recombination using the Gateway LR Clonase II Enzyme Mix (Invitrogen, Carlsbad, CA, USA). The pMDC32-*AtDjA3* vector was transferred into the *Agrobacterium tumefaciens* strain GV2260 by electroporation, and transformed into *Arabidopsis* plants WT (Col-0) by the "floral dip" method (Zhang et al., 2006). Transgenic lines carrying the *AtDjA3* gene (*35S::AtDjA3*) were selected on MS 0.5x medium containing 50 mg/mL hygromycin. To verify the expression levels of the *AtDjA3* gene in the two transgenic lines, RT-PCR analysis was carried out. For that, we used the following oligonucleotides: FwAtDjA3 5′-TGACGATGAAGATGATGACCATC-3′ and RvT-NOS 5′-ATTGCCAAATGTTTGAACGATCG-3′. As loading control, the *A. thaliana Actin8* (At1g49240) transcript was amplified using the FwAct8 5′- GCCAGTGGTCGTACAACCG-3′ and RvAct8 5′-CACGACCAGCAAGGTCGAGACG-3′ oligonucleotides. The T2 generation of transgenic plants was transferred into soil pots and grown in growth chambers under controlled conditions to produce seeds. Homozygous transgenic lines (T3) were used for the subsequent analysis of seed germination and in the stress tolerance assays.

Quantitative RT-PCR (qRT-PCR) of *AtDjA3* Gene Under Abiotic Stress

Total RNA from *A. thaliana* was obtained from seedlings as described above and used for qRT-PCR assays. Possible genomic DNA contamination was removed using DNase I (Invitrogen, Carlsbad, CA, USA). RNA concentration was measured in a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies) before and after treatment with DNase I. cDNA synthesis and qRT-PCR analysis were performed by the one-step assay using the Power SYBR® Green RNA-to-CT™ One-Step Kit (Applied Biosystems, USA). The expression levels of the *AtDjA3* gene under salt, osmotic, and exogenous ABA treatments were assessed in 15-day-old *A. thaliana* Col-0 plants. Seedlings were transferred to liquid MS 0.5x medium supplemented with NaCl (150 and 175 mM), glucose (5 and 6%), and ABA (1, 3, and 5 μM) for 12 and 24 h. Expression level analyses of the *AtDjA3* gene were performed using the following oligonucleotides: FwqRTJ3n 5′-TGACGATGAAGATGATGACCATC-3′ and RvqRTJ3n 5′-GCAAGAGACAAATTGGTTGGAG-3′ for the *AtDjA3* gene (At3g44110), and UBQ5-F 5′-TCGACGCTTCATCTCGTCCT-3′ and UBQ5-R 5′-CGCTGAACCTTTCCAGATCC-3′ for the *UBQ5* (At3g62250) control gene. The expression level of the *ABI3* gene was measured using RNAs obtained from 15-day-old Col-0, *Atdja3*-null mutant and *35S::AtDjA3*-8 overexpression lines that were previously treated with ABA (0 and 5 μM) for 12 and 24 h. Expression level analyses of the *ABI3* gene were performed using the following oligonucleotides: FwqABI3 5′-CACAGCCAGAGTTCCTTCCTT-3′ and RvqABI3 5′-ATGTGGCATGGGACCAGACT-3′ for the *ABI3* gene (At3g24650), and FwAPT1 5′-GTCATCCCCGACTTCCC TAA-3′ and RvAPT1 5′-AGGCATATCTGTTGTTGCAGGT-3′ for the *APT1* (At1g27450) control gene. cDNA synthesis and quantitative PCR analyses were done in a 10 μL reaction mixture containing 50 ng of total RNA as template using the Power SYBR® Green RNA-to-CT™ 1-Step Kit (Applied Biosystems) as described previously (Salas-Muñoz et al., 2012; Rodríguez-Hernández et al., 2014). For each sample, three biological replicates were analyzed with their respective technical replicates. Experiments were repeated at least twice and gave similar results.

Germination Assays Under Abiotic Stress and Hormone Treatments

Seeds of *A. thaliana* ecotype Col-0, *Atdja3*-null mutant line (*j3)* and *35S::AtDjA3* overexpression lines (*OvJ3*; T3) were germinated under different stress conditions. The effect of salt stress on germination was evaluated on MS 0.5x medium supplemented with 0, 125, and 150 mM NaCl. The effect of osmotic stress on germination was assessed on MS 0.5x medium without sucrose and supplemented with 0, 4, and 5% glucose. In addition, seeds were germinated in presence of different concentrations of ABA (0, 1, 3, and 5 μM). The ABA stock solution was prepared by dissolving ABA in small aliquots of 1N NaOH. The ABA stock was diluted with distilled water. The germination assays were carried out using 20 seeds per treatment. The seeds were germinated and grown vertically on petri dishes, and counted when the radicle emerged from the seed coat. In addition, the green cotyledon number was scored after 21 days of NaCl, glucose, or ABA treatments. Data are mean ± SE ($n = 20$) from five biological replicates. Experiments were repeated at least three times and gave similar results.

Seed Weight Estimation of Col-0, *Atdja3*-mutant, and *35S::AtDjA3* Over-Expression Lines

Seed weight was calculated from three replicates for each line ($n = 3$), where 500 seeds represent each replicate. Each seed lot (1-month post-harvest) was measured on an analytical scale, and weights are expressed in milligrams. Experiments were repeated at least three times and gave similar results.

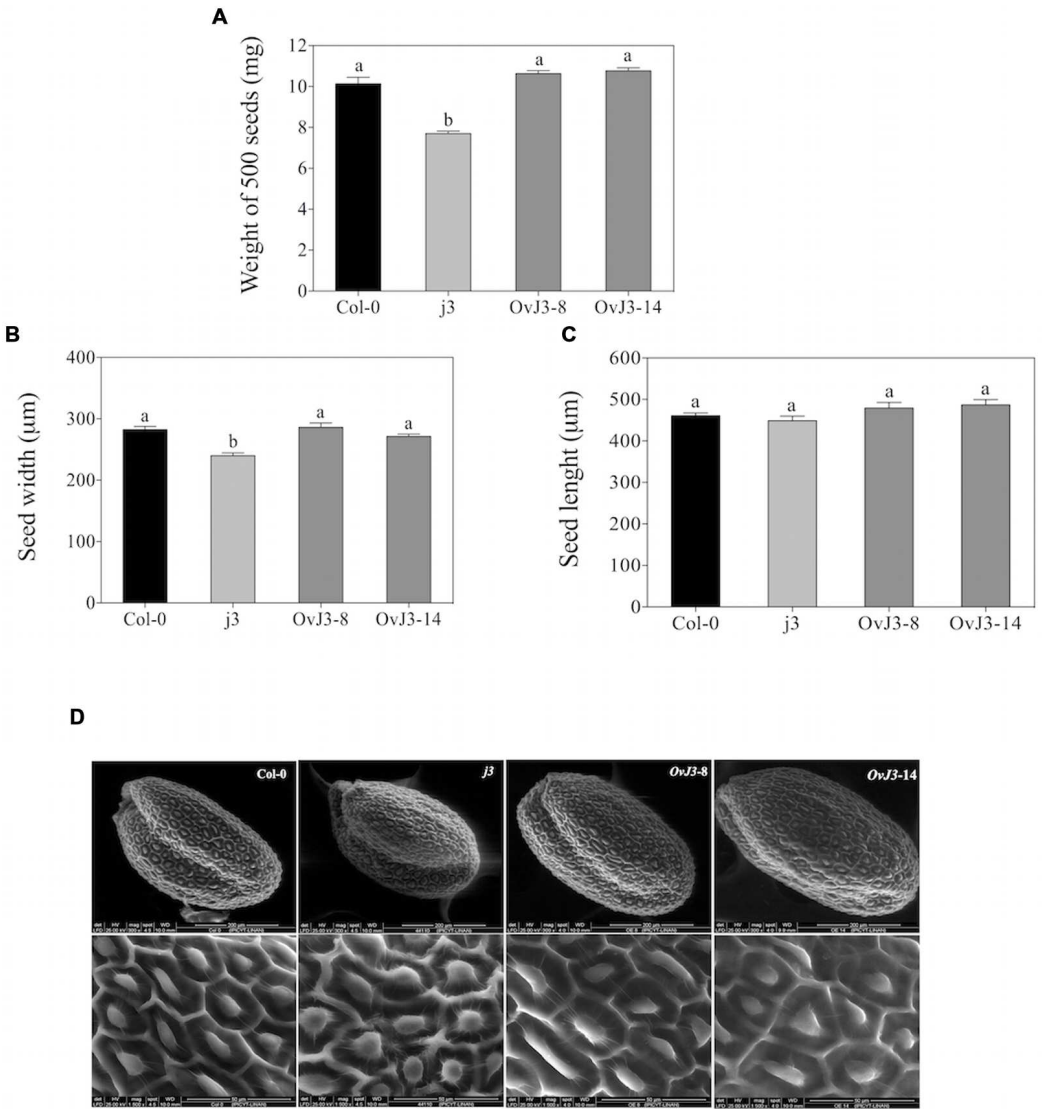


FIGURE 1 | Seed morphology of *Atdja3*-null mutant line (*j3*) and *35S::AtDjA3* overexpression lines (*OvJ3*-8 and -14). (A) Weight of seeds (mg) from the Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines. Data show means ± SE from the three groups of 500 dry seeds of each genotype. **(B)** Width and **(C)** Length of seeds (μm) from the Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines. **(D)** Scanning electron micrographs showing: whole seed, scale bar = 200 μm; seed coat, scale bar = 50 μm; of Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines. Error bars represent the means ± SE ($n = 10$). Different letters indicate significant differences between the Col-0, *Atdja3*-null mutant line, and *35S::AtDjA3* overexpression lines. One-way ANOVA analyzed significant differences among seeds lines through a Tukey's test ($p \leq 0.05$).

Microscopic Analysis by Environmental Scanning Electron Microscopy

For environmental scanning electron microscopy (eSEM) analysis, dried seeds were glued onto pure carbon-containing polymer films, and then fixed onto eSEM sample holders. The external seed morphology of Col-0, *Atdja3*-null mutant line (*j3*), and *35S::AtDjA3* overexpression lines (*OvJ3*) were evaluated. The seed width and length were measured with a high-resolution scanning electron microscope (eSEM/QUANTA 200 FEI, Low Vacuum/Water). Morphological seed assays, including width and length of seeds, were carried out using 10 seeds of each genotype. Photomicrographs were obtained with the eSEM in a pressure chamber at 90–100 Pa and voltages of 15.0 and 30.0 kV.

Statistical Analysis

To explore potential differences in germination, and green cotyledons among treatments for WT (Col-0), mutant line (*j3*) and transgenic lines (*35S::AtDjA3*-8 and -14), we used One-way ANOVA analysis through Tukey's multiple comparison post-test using GraphPad Software. The data are presented as the mean ± standard error. Differences at $p \leq 0.05$ were considered significant.

RESULTS

Seed Morphology of *Atdja3*-Null Mutant Line (*j3*) and *35S::AtDjA3* Overexpression Lines (*OvJ3*)

To address the biological functions of *AtDjA3* gene in seed morphology and response to abiotic stress, mutant, and overexpression lines were characterized. We selected the Salk_132923 line from the Salk T-DNA collection (Alonso et al., 2003), which contains a T-DNA insertion in the fourth exon of *AtDjA3* gene (Supplementary Figure S1A). The T-DNA homozygous line was identified by PCR, and the absence of the *AtDjA3* transcript was verified by RT-PCR (Supplementary Figure S1B), confirming that the Salk_132923 line is a null allele (*j3*) of the *AtDjA3* gene. In addition, we generated several transgenic *Arabidopsis* plants that overexpress the *AtDjA3* gene under the control of the CaMV 2X35S promoter (Supplementary Figure S1C). The expression levels of two independent *AtDjA3* overexpression lines (*35S::AtDjA3*-8 and -14) were determined by RT-PCR, observing expression of *AtDjA3* gene in all lines analyzed (*OvJ3*-8 and 14, respectively; Supplementary Figure S1D). Several parameters related to the seed morphology such as weight, length, width, and testa structure were analyzed. We observed that *j3* seeds showed a reduction in average seed weight (7.70 mg/500 seeds) in comparison to Col-0 (10.13 mg/500 seeds). With respect to *35S::AtDjA3*-8 and -14 overexpression lines (*OvJ3*-8 and -14), no significant differences in seed weight between overexpression lines and Col-0 were found (**Figure 1A**). In order to evaluate the seed width and length, and testa morphology of the *Atdja3*-null mutant line and *35S::AtDjA3* overexpression lines, micrographs of seeds were taken by eSEM (**Figures 1B–D**, respectively). In agreement with weight data, the *Atdja3*-null mutant seeds are reduced in width in comparison to Col-0 and *35S::AtDjA3* overexpression lines (**Figure 1B**). In addition, the surface of the seed testa in the *j3* mutant line displayed variations in the columella shape compared to Col-0 seeds (**Figure 1D**). With respect to the *35S::AtDjA3* transgenic lines, no significant differences between the overexpression lines and Col-0 were found in seed width and length (**Figures 1B,C**).

AtDjA3 Gene is Modulated Under Abiotic Stress

The expression patterns of the *AtDjA3* gene in response to salt and osmotic treatments, as well as the application to ABA hormone were assessed. qRT-PCR experiments were carried out in 15-day-old *A. thaliana* Col-0 seedlings subjected to NaCl (0, 150, and 175 mM), glucose (0, 5, and 6%), and ABA (0, 1, 3, and 5 μM) treatments for 12 and 24 h (**Figure 2**). In both salt treatments, an induction of the *AtDjA3* gene was observed at 12 and 24 h, except for the 24 h 150 mM NaCl treatment (**Figure 2A**). With respect to osmotic stress induction by glucose treatments, a slight expression of *AtDjA3* gene was observed at both treatment times, achieving the highest expression level with 6% glucose at 24 h (**Figure 2B**). Under ABA treatments, increases in *AtDjA3* gene expression were detected with 5 μM of hormone at 12 h, and 3 μM at 24 h (**Figure 2C**). The results showed that *AtDjA3* expression is modulated by abiotic stress.

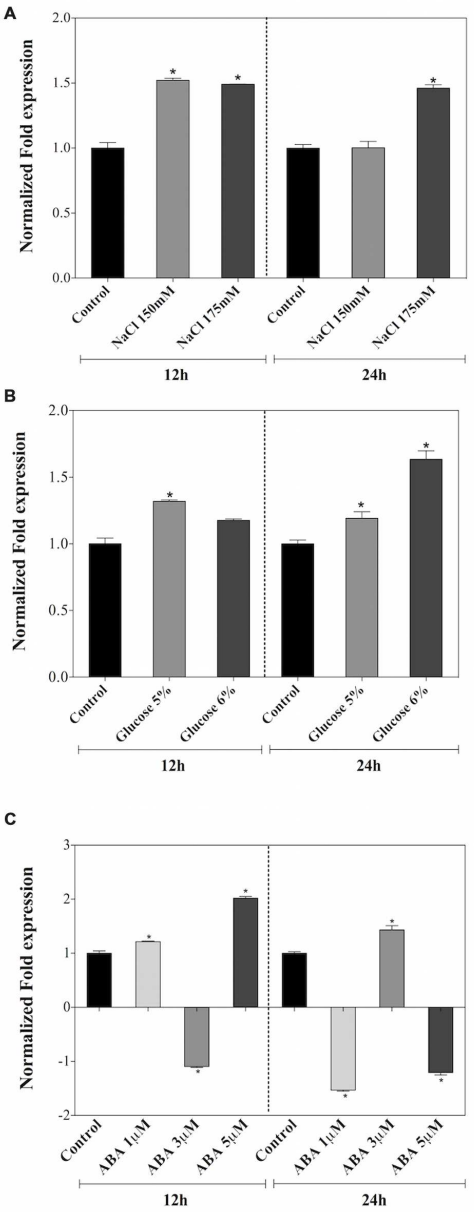


FIGURE 2 | Expression levels of *AtDjA3* gene under abiotic stress. The transcript level of *AtDjA3* in *A. thaliana* (Col-0) was determinate in 15-day-old seedlings grown on MS 0.5x liquid medium supplemented with 0, 150, and 175 mM NaCl **(A)**; 0, 5, and 6% glucose **(B)**; 0, 1, 3, and 5 μM ABA **(C)**. Gene expression was determined by qRT-PCR using SYBR green dye. Values represent fold change in expression level upon stressed seedlings compared to non-stressed control seedlings. Quantification was based on a cycle threshold value, with the expression level of the *AtDjA3* normalized to the *Arabidopsis UBQ5* gene. Bars represent mean ± SE ($n = 3$). In case of ratios lower than 1, the inverse of the ratio was estimated and the sign was changed. Asterisks indicate statistically significant differences between the samples treated and untreated, according to the One-way ANOVA analysis and multiple comparison Tukey's test ($p \leq 0.05$).

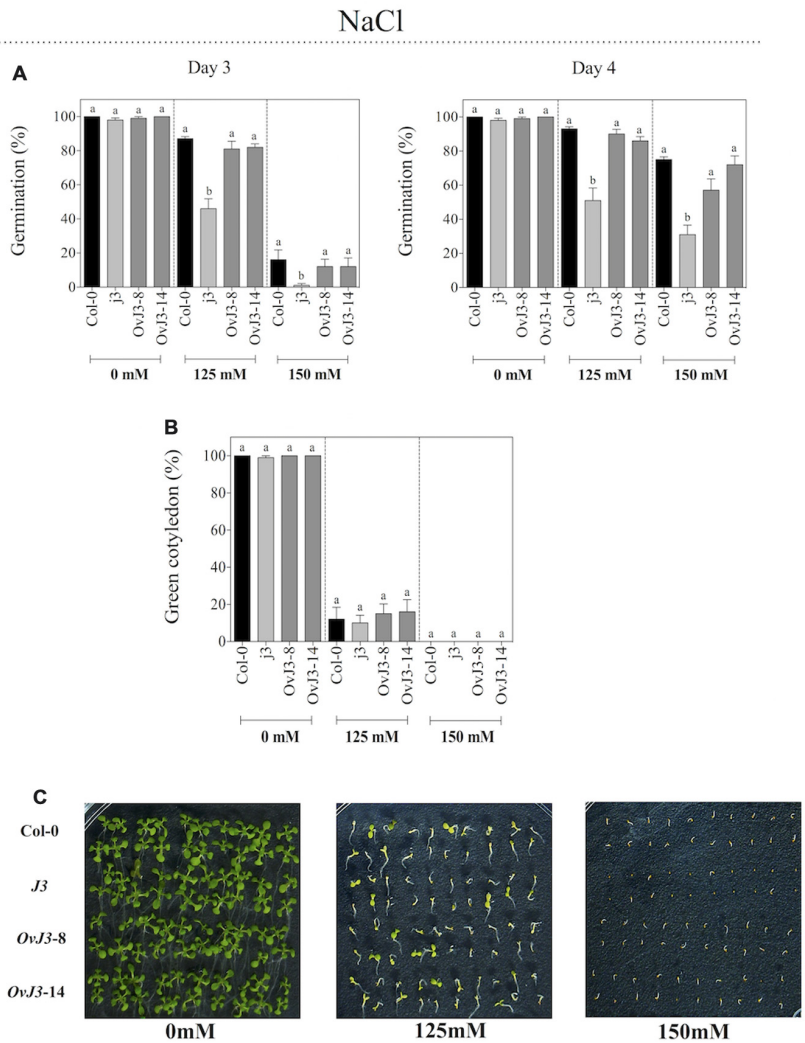


FIGURE 3 | Germination and green cotyledons development of *Arabidopsis* Col-0, *Atdja3*-null mutant line (*j3*), and *35S::AtDjA3* overexpression lines (*OvJ3*-8 and -14) under salinity. Seeds were germinated on MS 0.5x (1.5% sucrose) with 0, 125, and 150 mM NaCl. **(A)** Seed germination was evaluated at 3 and 4 days of treatment. **(B)** Percent of green cotyledons were recorded, and the phenotype at 21 days after the stress treatment is presented **(C)**. Bars represent the means ± SE ($n = 20$) of five replicates. Different letters indicate statistically significant differences between the Col-0, *Atdja3*-null mutant line, and *35S::AtDjA3* overexpression lines. One-way ANOVA was used to analyze the differences among treatments were explored through Tukey's test ($p \leq 0.05$).

The *j3* Mutant Line Showed Less Tolerance to Salt Stress in Germination and Post-germination Stages

To determine whether *AtDjA3* plays a role in *Arabidopsis* tolerance to salt stress, seeds of Col-0, *Atdja3*-null mutant line (*j3*) and *35S::AtDjA3* overexpression lines (*OvJ3*-8 and -14) were germinated on MS 0.5x medium containing 0, 125, and 150 mM NaCl. We observed that *j3* mutant line was affected during germination under salt treatments in comparison to Col-0 and the *35S::AtDjA3* overexpression lines (**Figure 3A**). At 4 days of 125 and 150 mM NaCl treatments there was a significant decrease in the germination rate of *j3* mutant line (51 and 31%, respectively) compared with Col-0 (93 and 75%, respectively). At 21 days after the germination under salt treatments (0, 125, and 150 mM), the percentage of green cotyledons was recorded (**Figures 3B,C**). At 125 mM NaCl, no significant differences in green cotyledons among the Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines were observed (**Figures 3B,C**). Under control conditions, no significant differences in germination rates and green cotyledons among the Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines were observed (**Figure 3**).

The *j3* Mutant Line Showed Increased Sensitivity to Glucose in Germination and Post-germination Growth

In addition to salt treatments, we analyzed the germination rates of Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines

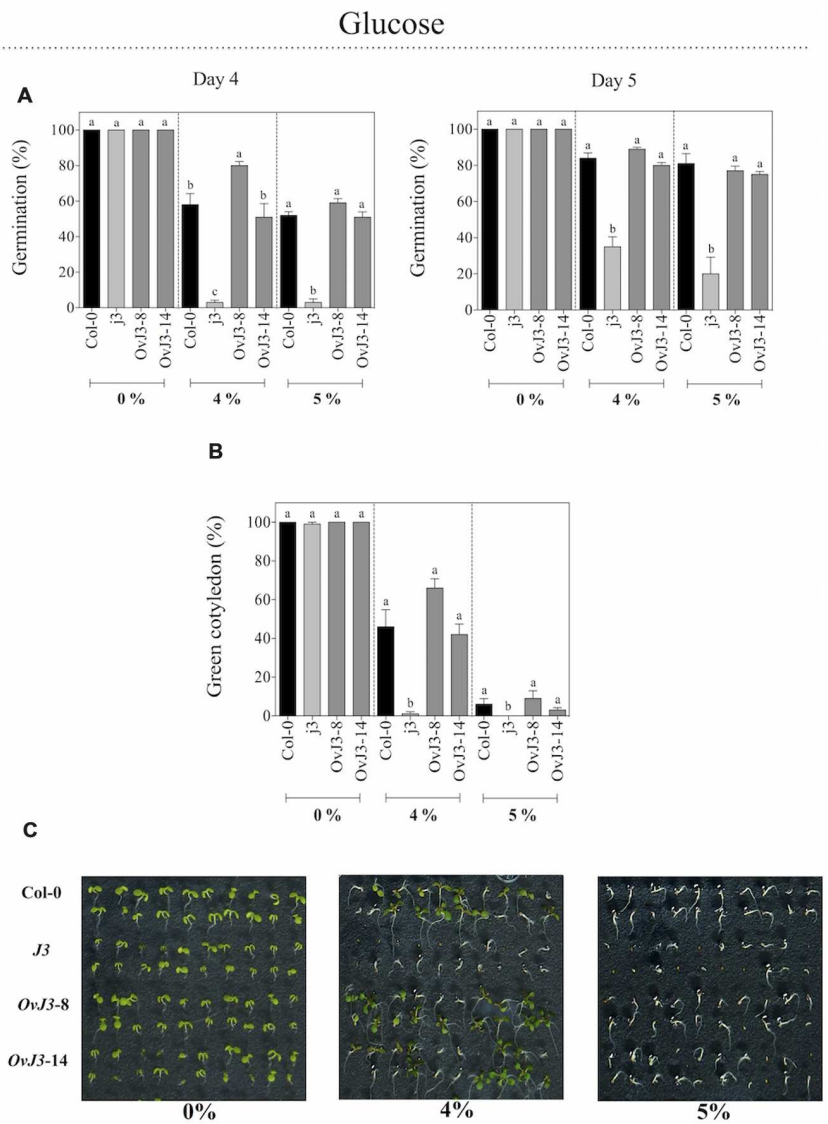


FIGURE 4 | Germination and green cotyledons development of *Arabidopsis* Col-0, *Atdja3*-null mutant line (*j3*), and *35S::AtDjA3* overexpression lines (*OvJ3*-8 and -14) under osmotic stress. Seeds were germinated on MS 0.5x (without sucrose) 0, 4, and 5% glucose **(A)** Seed germination was evaluated at 4 and 5 days of treatment. **(B)** Percent of green cotyledons were recorded, and the phenotype at 21 days after the stress treatment is presented **(C)**. Bars represent the means ± SE ($n = 20$) of five replicates. Different letters indicate statistically significant differences between the Col-0, *Atdja3*-null mutant line, and *35S::AtDjA3* overexpression lines. One-way ANOVA was used to analyze the differences among treatments were explored through Tukey's test ($p \leq 0.05$).

under osmotic stress, by sowing the seeds on MS 0.5x medium containing glucose at 0, 4, and 5% (**Figure 4**). When seeds were sown on glucose, the *j3* mutant line showed a clear reduction in its germination rate with respect to Col-0 and *35S::AtDjA3* overexpression lines (**Figure 4**). At 4 and 5 days, the osmotic sensitivity was noticed in the *j3* mutant line, which achieves only 3% of germination, whereas the Col-0 and overexpression lines exhibited percentages up to 50%. We analyzed the green cotyledons under glucose treatments (**Figures 4B,C**). As shown in germination, the *j3* mutant line showed an arrest in the development of the green cotyledons at both concentrations (**Figure 4C**).

The *j3* Mutant Line Showed Increased Sensitivity to ABA in Germination and Post-germination Growth

Inhibitory experiments of seed germination were carried out with Col-0, *j3* mutant line, and *35S::AtDjA3* overexpression lines on MS 0.5x medium containing 0, 1, 3, and 5 μM ABA (**Figure 5**). The *AtDjA3* gene disruption caused a germination sensitivity phenotype on ABA treatments. As observed at 3 and 4 days, the *j3* mutant line exhibited the lowest percentage of germination for all of the ABA concentrations assessed (**Figure 5A**). At the highest concentration of ABA (5 μM), the *j3* mutant line did

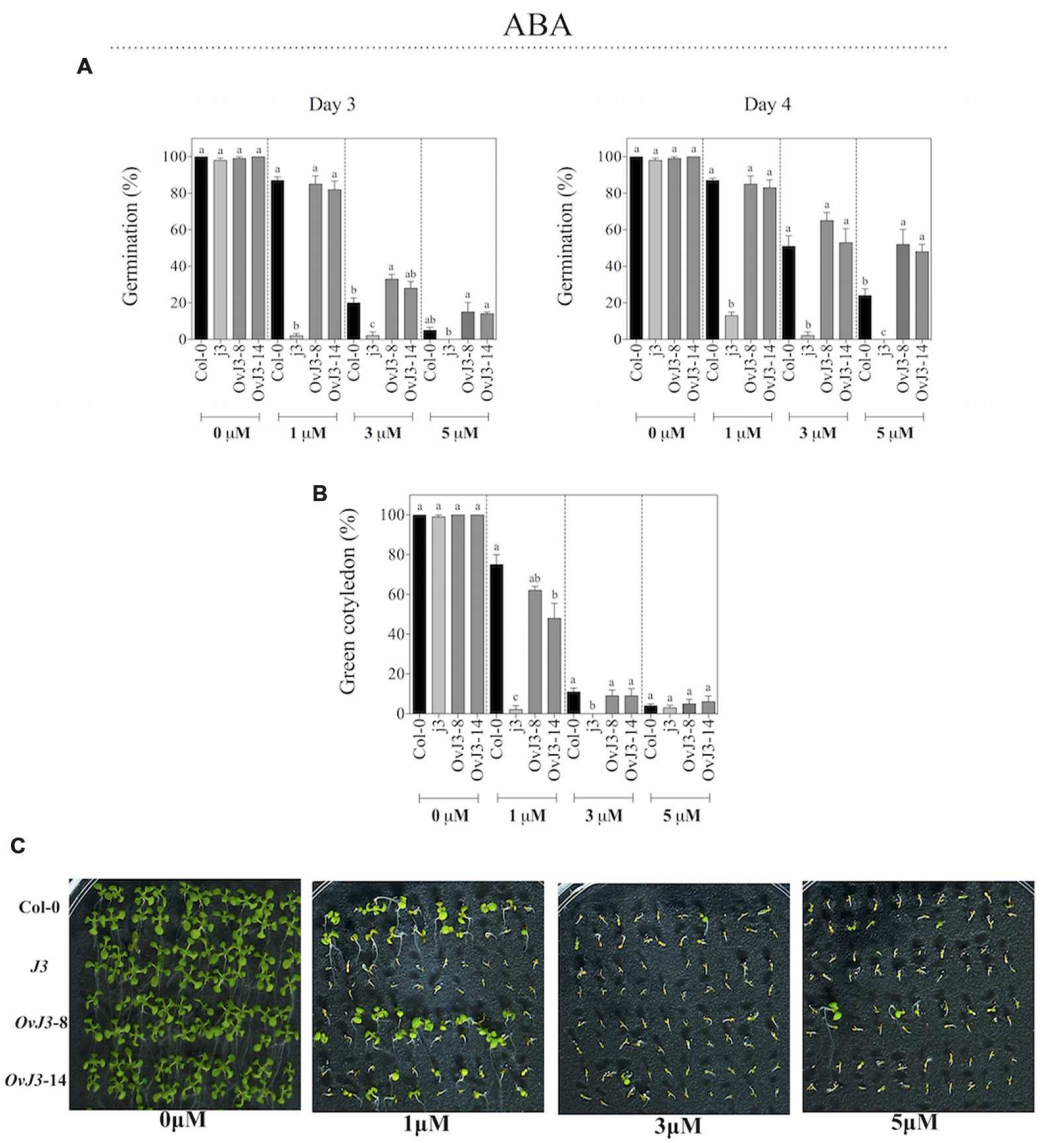


FIGURE 5 | ABA effect on germination and green cotyledons of *Arabidopsis* Col-0, *Atdja3*-null mutant line (*j3*), and *35S::AtDjA3* overexpression lines (*OvJ3*-8 and -14). Seeds were germinated on MS 0.5x (1.5% sucrose) with 0, 1, 3, and 5 µM ABA. **(A)** Seed germination was evaluated at 3 and 4 days of treatment. **(B)** Percent of green cotyledons were recorded, and the phenotype at 21 days after the stress treatment is presented **(C)**. Bars represent the means ± SE ($n = 20$) of five replicates. Different letters indicate statistically significant differences between the Col-0, *Atdja3*-null mutant line, and *35S::AtDjA3* overexpression lines. One-way ANOVA was used to analyze the differences among treatments were explored through Tukey's test ($p \leq 0.05$).

not germinate, whereas the Col-0 achieved a 24% germination, and the transgenic lines (*35S::AtDjA3*) exhibited more than 40% germination after 4 days (**Figure 5A**). At 21 days after the germination under ABA treatment (0, 1, 3, and 5 µM), the percentage of green cotyledons was recorded (**Figures 5B,C**). The *j3* mutant line exhibits a lower percentage of green cotyledons under ABA treatments. At 1 µM ABA, the *j3* mutant line showed post-germination growth arrest (only 2% of green cotyledons), whereas the Col-0 and the overexpression lines achieved more than 40% of green cotyledons. These data revealed that the *j3* mutant line exhibited an ABA-sensitive phenotype.

Abscisic Acid-Insensitive 3 (*ABI3*) Gene Expression in the *Atdja3*-Null Mutant Line (*j3*) and *35S::AtDjA3* Overexpression Lines (*OvJ3*) Under ABA Treatment

Based on ABA sensitivity observed in the *j3* mutant line, we examined the expression of *ABI3* transcription factor. For this, qRT-PCR expression analyses were carried out in 15-day-old plants of Col-0, *j3* mutant line, and *35S::AtDjA3-8* overexpression lines by being subjected to ABA treatment (5 µM) for 12 and 24 h (**Figure 6**). The *j3* mutant line showed higher *ABI3*

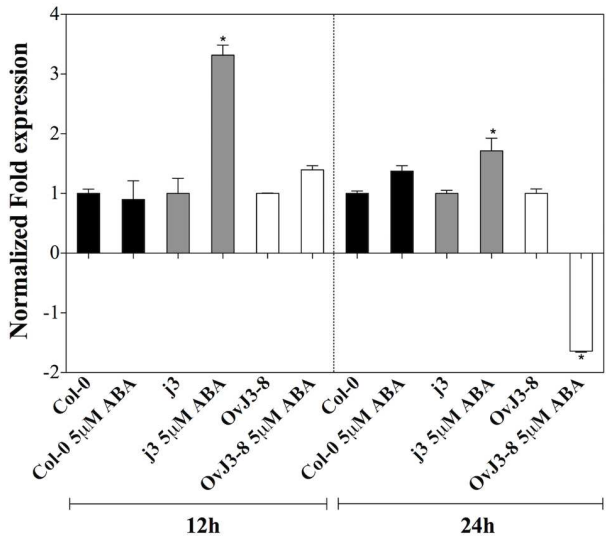


FIGURE 6 | *ABI3* expression levels on *Arabidopsis* Col-0, *Atdja3*-null mutant line (*j3*), and *35S::AtDjA3-8* over-expression line under ABA treatment. Fifteen-day-old *Arabidopsis* seedlings were placed on MS 0.5x liquid medium supplemented with 0 and 5 μM ABA at 12 and 24 h. The expression levels of *ABI3* were determined by qRT-PCR using SYBR green dye. Data are expressed as relative mRNA level compared to the plants of each line without ABA treatments, and were calculated after normalization to the *Arabidopsis APT1* gene using the comparative threshold method. In case of ratios lower than 1, the inverse of the ratio was estimated and the sign was changed. Bars represent mean ± SE ($n = 3$). Asterisks indicate statistically significant differences between the samples treated and untreated, according to the One-way ANOVA analysis and multiple comparison Tukey's test ($p \leq 0.05$).

transcript levels under ABA treatment during the entire time course analyzed compared to the parental (Col-0). The highest expression of the *ABI3* gene in the mutant line was observed at 12 h (**Figure 6**). In contrast, the *ABI3* gene was repressed in the *35S::AtDjA3-8* overexpression line at 24 h under ABA treatment.

DISCUSSION

One of the major molecular mechanisms to re-establish cellular homeostasis and protect cellular components under abiotic stress is the expression of stress response genes, which encode molecular chaperones such as HSP. HSP40, also known as J-proteins, act as molecular chaperones, and are involved in many cellular processes, including development, signal transduction, and resistance to environmental stresses (Rajan and D'Silva, 2009). We have characterized the *Arabidopsis thaliana AtDjA3* gene (At3g44110), which encodes a J-protein, belonging to group I (also known as group A). To determine the possible functional roles of *AtDjA3* in response to abiotic stresses, a T-DNA mutant line and overexpression lines (*35S::AtDjA3*) were obtained and characterized. Remarkably, we found that the *Atdja3*-null mutant line (*j3*) produced smaller and lighter seeds in comparison to parental Col-0. On the other hand, seeds from the overexpression lines showed no significant differences in seed weight or size compared to Col-0 seeds. We noticed that the lack of *AtDjA3* transcript altered the columella shape of the seeds; these alterations could be due to alterations in the shape and size of seed. In agreement with our proposal, the *Arabidopsis* microarray database (*Arabidopsis* eFP Browser[2]) reports that *AtDjA3* gene is induced during stages 8, 9, and 10 of seed development, and also in dry seed (Schmid et al., 2005). Thus, the loss of function of *AtDjA3* gene could be an important factor in seed formation.

We have showed that abiotic stressors modulated the *AtDjA3* gene, including NaCl, glucose, and the application of the ABA hormone. The highest accumulation of *AtDjA3* transcript was observed under ABA treatment, in particular with 5 μM ABA at 12 h. Characterization of the *AtDjA3* gene in the germination process of *Arabidopsis* seeds under abiotic stress conditions using the *j3* mutant line and *35S::AtDjA3* overexpression lines revealed that the loss-of-function of *AtDjA3* resulted in seeds with sensitivity to salt and osmotic stresses. Furthermore, seeds of the *j3* mutant line exhibited increased sensitivity to ABA during germination compared to the parental Col-0. Conversely, *35S::AtDjA3* overexpression lines showed the highest rate of germination after 4 days at the maximum concentration of ABA (5 μM). We also examined the effects of the *j3* mutant line on the post-germination growth in the salt and osmotic stress response. The results showed that cotyledon development in *j3* mutant line was severely inhibited during glucose and ABA treatments, while during salt treatments no differences were observed. On the other hand, the *AtDjA3* overexpression lines showed a similar behavior to parental Col-0 in germination and cotyledon greening under salt and osmotic treatments.

A total of 120 J-domain proteins have been identified in the *Arabidopsis* genome, which represents a large and diverse family of molecular chaperones (Rajan and D'Silva, 2009). Although their functions are mostly uncharacterized, J-proteins have been implicated in plant stress response. In particular, the role of *AtDjA3* in heat and salt stress has been documented. For instance, Li et al. (2007) reported that *AtDjA3* and its paralogous gene, *AtDjA2*, improve *Arabidopsis* thermotolerance. In addition, Yang et al. (2010) showed that plants lacking *AtDjA3* gene are more sensitive to salt at alkaline pH, and exhibit decrease plasma membrane H^+-ATPase activity. The authors reported that under alkaline conditions, *AtDjA3* interacts with protein kinase 5 (PKS5), repressing PKS5 kinase activity to release plasma membrane H^+-ATPase. These reports are consistent with our findings, showing that the chaperone AtDjA3 plays a key role during abiotic stress tolerance. Moreover, others J-proteins have been detected during salt and heat stress. For instance, *ANJ1*, a DnaJ gene from *Atriplex nummularia*, was induced by heat and salt stress (Zhu et al., 1993). Similarly, the expression of *SGJ3* (DnaJ-like) was rapidly induced in Japanese willow (*Salix gilgiana* S.) plants upon exposure to heat and salt stress (Futamura et al., 1999). Overexpression of *Arabidopsis DnaJ* gene (type I,

[2]www.bar.utoronto.ca

At2g22360) in *E. coli* and *Arabidopsis* plants exhibited increased tolerance to NaCl stress (Zhichang et al., 2010).

We noticed that the disruption of *AtDjA3* gene resulted in ABA hypersensitivity. ABA plays an important role in developmental processes such as seed maturation, including synthesis of seed storage proteins and lipids, seed desiccation tolerance, dormancy, control of germination, and the subsequent commitment to seedling growth, and adaptive responses to environmental stimuli in plants (Finkelstein et al., 2002; Cutler et al., 2010; Hubbard et al., 2010; Raghavendra et al., 2010; Fujita et al., 2011). The *j3* mutant line showed sensitivity to ABA in germination and post-germination stages in comparison to parental Col-0. Conversely, *35S::AtDjA3* overexpression lines had less sensitivity to 5 μM ABA in germination than those of Col-0. We evaluated the *ABA-insensitive 3* (*ABI3*) gene expression in the Col-0, *j3* and *35S::AtDjA3* overexpression lines under ABA treatment. We found that both the lack of *AtDjA3* transcript and its constitutive overexpression altered *ABI3* gene expression. *ABI3* expression was induced in the *j3* mutant line in comparison to WT (Col-0) at 12 and 24 h after ABA treatment. This increased induction of *ABI3* transcript in the *j3* mutant line, a key factor in ABA signaling, could be correlated with the ABA hypersensitivity phenotype observed in the *j3* mutant line during germination and post-germination growth. The transcriptional factor ABI3 is considered to be essential for the regulation of seed specific development, so this factor determines the ABA sensitivity and plays a key role in desiccation tolerance and dormancy during zygotic embryogenesis (Zhang et al., 2005). *ABI3* transcript and protein levels are abundant in maturing and mature seeds, but disappear soon after germination. However, these levels can be modulated by ABA or osmotic stress during the time period when post-germination growth arrest occurs (Lopez-Molina et al., 2001, 2002). In contrast with the *j3* mutant, we found that transgenic plants overexpressing *AtDjA3* were slightly more resistant to 5 μM ABA compared to the WT during germination. This could be explained by the down-regulation of the *ABI3* transcript at 5 μM ABA treatment in *35S::AtDjA3* overexpressing line, which is contrasting behavior to that observed in the *Atdja3*-null mutant background.

The results presented here showed that the *j3* mutant line generates smaller seeds, which were more sensitive to abiotic stress and exogenous application of ABA. This phenotype of *j3* mutant line to stress treatments reveals that *AtDjA3* gene might have an important role during the germination process, and provides new insights into abiotic stress responses mediated by chaperones.

AUTHOR CONTRIBUTIONS

SS-M, AR-H, MO-A, FS-B designed and carried out the experiments, analyzed the results, and wrote the manuscript. JJ-B designed the research, contributed scientific advice, correction, wrote and revision of the manuscript. All authors have read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by the CONACYT: Investigacioìn Ciencia Baìsica CB-2013-221075, and Fortalecimiento de infraestructura para la consolidación INFR-2014-01-224800 funds. We are grateful to Dr. Steffen Graether for a grammatical review, and M. C. Alicia Becerra Flora for their technical assistance. We thank M. C. Gladis Delgado Labrada for their technical assistance in the Environmental scanning electron microscopy (eSEM) at LINAN, IPICYT.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00220

REFERENCES

Alonso, J. M., Stepanova, A. N., and Leisse, T. J. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653–657. doi: 10.1126/science.1086391

Bukau, B., and Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* 92, 351–366.

Craig, E. A., Huang, P., Aron, R., and Andrew, A. (2006). The diverse roles of J-proteins, the obligate Hsp70 co-chaperone. *Rev. Physiol. Biochem. Pharmacol.* 156, 1–21.

Curtis, M. D., and Grossniklaus, U. (2003). A gateway cloning vector set for high throughput functional analysis of genes in planta. *Plant Physiol.* 133, 462–469. doi: 10.1104/pp.103.027979

Cutler, S. R., Rodriguez, P. L., Finkelstein, R. R., and Abrams, S. R. (2010). Abscisic acid: emergence of a core signaling network. *Annu. Rev. Plant Biol.* 61, 651–679. doi: 10.1146/annurev-arplant-042809-112122

Daszkowska-Golec, A. (2011). *Arabidopsis* seed germination under abiotic stress as a concert of action of phytohormones. *OMICS* 15, 763–774. doi: 10.1089/omi.2011.0082

Finkelstein, R. R., Gampala, S. S., and Rock, C. D. (2002). Abscisic acid signalling in seeds and seedlings. *Plant Cell* 14, S15–S45.

Fujita, Y., Fujita, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J. Plant Res.* 124, 509–525. doi: 10.1007/s10265-011-0412-3

Futamura, N., Ishii-Minami, N., Hayashida, N., and Shinohara, K. (1999). Expression of DnaJ homologs and Hsp70 in the Japanese willow (*Salix gilgiana* Seemen). *Plant Cell Physiol.* 40, 524–531. doi: 10.1093/oxfordjournals.pcp.a029573

Hubbard, K. E., Nishimura, N., Hitomi, K., Getzoff, E. D., and Schroeder, J. I. (2010). Early abscisic acid signal transduction mechanisms: newly discovered components and newly emerging questions. *Genes Dev.* 24, 1695–1708. doi: 10.1101/gad.1953910

Koornneef, M., Bentsink, L., and Hilhorst, H. (2002). Seed dormancy and germination. *Curr. Opin. Plant Biol.* 5, 33–36. doi: 10.1016/S1369-5266(01)00219-9

Li, G. L., Chang, H., Li, B., Zhou, W., Sun, D. Y., and Zhou, R. G. (2007). The roles of the atDjA2 and atDjA3 molecular chaperone proteins in improving thermotolerance of *Arabidopsis thaliana* seedlings. *Plant Sci.* 173, 408–416. doi: 10.1016/j.plantsci.2007.06.013

Li, G. L., Li, B., Liu, H. T., and Zhou, R. G. (2005). The responses of AtJ2 and AtJ3 gene expression to environmental stresses in *Arabidopsis*. *J. Plant Physiol. Mol. Biol.* 31, 47–52.

Lopez-Molina, L., Mongrand, S., and Chua, N. H. (2001). A post-germination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4782–4787. doi: 10.1073/pnas.081594298

Lopez-Molina, L., Mongrand, S., McLachlin, D. T., Chait, B. T., and Chua, N. H. (2002). ABI5 acts downstream of ABI3 to execute an ABA- dependent growth arrest during germination. *Plant J.* 32, 317–328. doi: 10.1046/j.1365-313X.2002.01430.x

Lu, Z., and Cyr, D. M. (1998). The conserved carboxyl terminus and zinc finger-like domain of the co-chaperone Ydj1 assist Hsp70 in protein folding. *J. Biol. Chem.* 273, 5970–5978. doi: 10.1074/jbc.273.10.5970

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi: 10.1093/nar/8.19.4321

Raghavendra, A. S., Gonugunta, V. K., Christmann, A., and Grill, E. (2010). ABA perception and signaling. *Plant Sci.* 15, 395–401. doi: 10.1016/j.tplants.2010.04.006

Rajan, V. B. V., and D'Silva, P. (2009). *Arabidopsis thaliana* J-class heat shock proteins: cellular stress sensors. *Funct. Integr. Genomics* 9, 433–446. doi: 10.1007/s10142-009-0132-0

Rao, K. M., Raghavendra, A. S., and Reddy, K. J. (eds). 2006). *Physiology and Molecular Biology of Stress Tolerance in Plants*. Dordrecht: Springer.

Rodríguez-Hernández, A. A., Ortega-Amaro, M. A., Delgado-Sánchez, P., Salinas, J., and Jiménez-Bremont, J. F. (2014). AtGRDP1 gene encoding a glycine-rich domain protein is involved in germination and responds to aba signaling. *Plant Mol. Biol. Rep.* 32, 1187–1202. doi: 10.1007/s11105-014-0714-4

Salas-Muñoz, S., Gómez-Anduro, G., Delgado-Sánchez, P., Rodríguez-Kessler, M., and Jiménez-Bremont, J. F. (2012). The *Opuntia streptacantha* OpsHSP18 gene confers salt and osmotic stress tolerance in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 13, 10154–10175. doi: 10.3390/ijms130810154

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi: 10.1038/ng1543

Shi, Y. Y., Hong, X. G., and Wang, C. C. (2005). The C-terminal (331-376) sequence of *Escherichia coli* DnaJ is essential for dimerization and chaperone activity: a small angle X-ray scattering study in solution. *J. Biol. Chem.* 280, 22761–22768. doi: 10.1074/jbc.M503643200

Vallejo, A. J., Yanovsky, M. J., and Botto, J. F. (2010). Germination variation in *Arabidopsis thaliana* accessions under moderate osmotic and salt stresses. *Ann. Bot.* 106, 833–842. doi: 10.1093/aob/mcq179

Vierling, E. (1991). The roles of heat shock proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42, 579–620. doi: 10.1146/annurev.pp.42.060191.003051

Wall, D., Zylicz, M., and Georgopoulos, C. (1994). The NH2-terminal 108 amino acids of the *Escherichia coli* DnaJ protein stimulate the ATPase activity of DnaK and are sufficient for lambda replication. *J. Biol. Chem.* 269, 5446–5451.

Walsh, P., Bursac, D., Law, Y., Cyr, D., and Lithgow, T. (2004). The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep.* 5, 567–571. doi: 10.1038/sj.embor.7400172

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9, 244–252. doi: 10.1016/j.tplants.2004.03.006

Wehmeyer, N., and Vierling, E. (2000). The expression of small heat shock proteins in seeds responds to discrete developmental signals and suggests a general protective role in desiccation tolerance. *Plant Physiol.* 122, 1099–1108. doi: 10.1104/pp.122.4.1099

Xue, Y., Peng, R., Xiong, A., Li, X., Zha, D., and Yao, Q. (2010). Over-expression of heat shock protein gene hsp26 in *Arabidopsis thaliana* enhances heat tolerance. *Biol. Plant.* 54, 105–111. doi: 10.1016/j.jplph.2008.11.013

Yang, Y., Qin, Y., Xie, C., Zhao, F., Zhao, J., Liu, D., et al. (2010). The *Arabidopsis* chaperone J3 regulates the plasma membrane H+-ATPase through interaction with the PKS5 Kinase. *Plant Cell* 22, 1313–1332. doi: 10.1105/tpc.109.069609

Zhang, X., Garreton, V., and Chua, N. H. (2005). The AIP2 E3 ligase acts as a novel negative regulator of ABA signaling by promoting ABI3 degradation. *Genes Dev.* 19, 1532–1543. doi: 10.1101/gad.1318705

Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W., and Chua, N. H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Prot.* 1, 641–646. doi: 10.1038/nprot.2006.97

Zhichang, Z., Wanrong, Z., Jinping, Y., Jianjun, Z., Xufengand, L. Z. L., and Yang, Y. (2010). Overexpression of *Arabidopsis* DnaJ (Hsp40) contributes to NaCl-stress tolerance. *Afr. J. Biotechnol.* 9, 972–978.

Zhu, J. K., Shi, J., Bressan, R. A., and Hasegawa, P. M. (1993). Expression of an *Atriplex nummularia* gene encoding a protein homologous to the bacterial molecular chaperone DnaJ. *Plant Cell* 5, 341–349. doi: 10.2307/3869601

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Salas-Muñoz, Rodríguez-Hernández, Ortega-Amaro, Salazar-Badillo and Jiménez-Bremont. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 31 August 2016
doi: 10.3389/fpls.2016.01325

Genome-Wide Analysis of the Glutathione *S*-Transferase Gene Family in *Capsella rubella*: Identification, Expression, and Biochemical Functions

*Gang He[1,2†], Chao-Nan Guan[3†], Qiang-Xin Chen[1], Xiao-Jun Gou[2], Wei Liu[2], Qing-Yin Zeng[1] and Ting Lan[1]**

[1] Functional Genomics and Protein Evolution Group, State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing, China, [2] The Key Laboratory of Medicinal and Edible Plants Resources Development of Sichuan Education Commission, Chengdu University, Chengdu, China, [3] College of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing, China

Extensive subfunctionalization might explain why so many genes have been maintained after gene duplication, which provides the engine for gene family expansion. However, it is still a particular challenge to trace the evolutionary dynamics and features of functional divergences in a supergene family over the course of evolution. In this study, we identified 49 Glutathione *S*-transferase (GST) genes from the *Capsella rubella*, a close relative of *Arabidopsis thaliana* and a member of the mustard family. *Capsella* GSTs can be categorized into eight classes, with tau and phi GSTs being the most numerous. The expansion of the two classes mainly occurs through tandem gene duplication, which results in tandem-arrayed gene clusters on chromosomes. By integrating phylogenetic analysis, expression patterns, and biochemical functions of *Capsella* and *Arabidopsis* GSTs, functional divergence, both in gene expression and enzymatic properties, were clearly observed in paralogous gene pairs in *Capsella* (even the most recent duplicates), and orthologous GSTs in *Arabidopsis*/*Capsella*. This study provides functional evidence for the expansion and organization of a large gene family in closely related species.

Keywords: gene family, gene duplication, genome, enzyme activity, functional divergence

OPEN ACCESS

Edited by:
Juan Francisco Jimenez Bremont, Instituto Potosino de Investigacion Cientifica y Tecnologica, Mexico

Reviewed by:
Chunyu Zhang, Huazhong Agricultural University, China
Pablo Peláez, National Autonomous University of Mexico, Mexico
Abraham Cruz-Mendívil, Instituto Politécnico Nacional, Mexico

***Correspondence:**
Ting Lan
lanting@ibcas.ac.cn

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *18 April 2016*
Accepted: *18 August 2016*
Published: *31 August 2016*

Citation:
He G, Guan C-N, Chen Q-X, Gou X-J, Liu W, Zeng Q-Y and Lan T (2016) Genome-Wide Analysis of the Glutathione S-Transferase Gene Family in Capsella rubella: Identification, Expression, and Biochemical Functions. Front. Plant Sci. 7:1325. doi: 10.3389/fpls.2016.01325

INTRODUCTION

Glutathione *S*-transferases (GSTs; EC 2.5.1.18) are multifunctional proteins encoded by a large gene family that is found in most organisms. As classical phase II detoxification enzymes, GSTs mainly catalyze the conjugation of reduced glutathione (GSH) with a wide variety of reactive electrophiles (Hayes et al., 2005). In plants, GST proteins are involved in several crucial physiological and developmental processes, including xenobiotic (e.g., herbicides) detoxification, signal transduction, isomerization, and protection against oxidative damages, UV radiation, and heavy metal toxins (Dixon et al., 2010; Cummins et al., 2011). Based on amino acid sequence similarity and gene organization, plant GSTs have been categorized into eight classes: phi, tau, theta, zeta, lambda, dehydroascorbate reductase (DHAR), tetrachlorohydroquinone dehalogenase (TCHQD) and the

class containing the γ-subunit of the eukaryotic translation elongation factor 1B (EF1Bγ) (Oakley, 2005; Lan et al., 2009; Dixon and Edwards, 2010a). We recently identified two new GST classes (hemerythrin and iota) in non-vascular plants (Liu et al., 2013). Among the ten GST classes, phi, tau, lambda, and DHAR GSTs are considered unique to plants (Frova, 2006).

In plants, tau and phi class GSTs are the most numerous and play important roles in detoxification of xenobiotics (Frova, 2003). Overexpression of tau or phi GSTs in plants can increase tolerance to oxidation, herbicides, salinity, and chilling (Roxas et al., 1997; Cummins et al., 1999; Karavangeli et al., 2005; Benekos et al., 2010; Sharma et al., 2014). These proteins also participate in non-catalytic functions, e.g., binding/transport and signaling (Marrs, 1996; Lieberherr et al., 2003; Kitamura et al., 2004). Lambda and DHAR GSTs do not exhibit activity toward xenobiotics but are considered to be involved in redox and thiol transfer reactions (Dixon et al., 2002a; Dixon and Edwards, 2010b). DHAR GSTs have key functions not only in the ascorbate-GSH recycling reaction but also in stress resistance (Kwon et al., 2003; Chen and Gallie, 2006; Ushimaru et al., 2006). Recent studies demonstrated that some stress-inducible lambda GSTs could selectively bind flavonols and serve as antioxidants (Dixon and Edwards, 2010b; Dixon et al., 2011). The theta and zeta GSTs have counterparts in the mammalian system and function mainly as GSH-dependent peroxidases and isomerases (Thom et al., 2001; Basantani and Srivastava, 2007). GSTs in EF1Bγ class contain two domains: a typical GST domain and an EF1Bγ domain. The GST domain of EF1Bγ class GSTs functions as GSH peroxidases (Vickers et al., 2004).

Capsella rubella is from the same family as *Arabidopsis thaliana*. *C. rubella* is a model species widely used for studying natural variation in adaptive traits, such as flowering time (Guo et al., 2012). This species is also a good model for understanding the evolution of self-fertilization (Guo et al., 2009). In *Arabidopsis*, the haploid set consists of five chromosomes, whereas its close relative *C. rubella* has $n = 8$ chromosomes (Boivin et al., 2004). The progenitors of the lineage leading to *A. thaliana* and *C. rubella* diverged approximately 10 million years ago (Acarkan et al., 2000; Koch and Kiefer, 2005). The *C. rubella* genome has been completely sequenced (Slotte et al., 2013), thus facilitating the understanding of the evolutionary relationship between *C. rubella* and its relative *A. thaliana* from the gene family level. In this study, we performed genome-wide annotation of the GST gene family of *C. rubella*. Through phylogenetic analysis with expression and functional assays, we provided detailed characterization of the organization, gene expression pattern, and enzymatic properties of the GST members. Extensive functional divergence was observed among members within tandem-arrayed GST clusters and between paralogous gene pairs. Through comparative analyses of this family in *C. rubella* and *A. thaliana*, we examined the lineage-specific loss/gain events, and divergences in expression and substrate specificity in the orthologous GSTs. The genome-wide, multifaceted approach we employed provides new insights into the process of gene family evolution between closely related species.

MATERIALS AND METHODS

Gene Identification and Nomenclature

To identify putative GST members in *C. rubella*, we performed TBLASTN searches with default algorithm parameters in the *Capsella* genome database, version 1.0[1], using 55 GST protein sequences of *Arabidopsis* (Dixon and Edwards, 2010a), 81 of *populus* (Lan et al., 2009), and 575 of other plants, animals, fungi, and bacteria (Supplementary Table S5) as queries. These 575 full-length GSTs represent 36 GST sub-families defined by the NCBI Conserved Domain Database (CDD; Marchler-Bauer et al., 2011). All potential candidates identified were examined using the Pfam[2] and CDD[3] database to confirm the presence of typical GST N- and C-terminal domains in their protein structures. Preliminary classification of GST genes into subfamilies was performed using phylogenetic analysis. The proteins, which clustered with soluble cytosolic GSTs, have an ancient monophyletic origin (Dixon and Edwards, 2010a). They were used in subsequent analyses. Next, *Capsella* GSTs were amplified from genomic DNA and mRNA from mixed tissues of *C. rubella*, cloned into the pGEM-T Easy Vector (Promega), and sequenced in both directions to verify the gene sequences. The primers used for gene amplification are listed in Supplementary Table S2. Complete manual curation of the gene sequences and structures based on expressed sequence tag (EST) databases and experimental support was further performed to rectify incorrect start codon predictions, splicing errors, missed or extra exons, and incorrectly predicted pseudogenes. For genes that went undetected by PCR (5 out of 49 in this study), their gene structures were assumed to be identical to those of their closest phylogenetic relatives. This approach was adapted from other studies (Meyers et al., 2003).

The nomenclature for *Capsella* GSTs follows the system suggested by Dixon et al. (2002b) for plant GSTs. A univocal name was assigned to each *Capsella* GST gene consisting of two italic letters *Cr* denoting the source organism, the family name (e.g., *CrGSTU*, *CrGSTF*, *CrGSTT*, *CrGSTZ*, *CrGSTL*, *CrTCHQD*, *CrDHAR*, and *CrEF1Bγ* corresponding to tau, phi, theta, zeta, lambda, TCHQD, DHAR, and EF1Bγ classes, respectively) and a progressive number for each gene (e.g., *CrGSTU1*).

Phylogenetic Analyses

Full-length amino acid sequences were aligned using MUSCLE software[4] and adjusted manually with BioEdit (Hall, 1999). Phylogenetic analysis was performed using the maximum-likelihood (ML) method in PHYML software (Guindon and Gascuel, 2003) with the Jones, Taylor, and Thornton (JTT) amino acid substitution model. GRX2 protein from *Escherichia coli* was chosen as an out-group during phylogenetic analysis of the *Capsella* GST family, as cytosolic GSTs are thought to be derived from the GRX2 (Holm et al., 2006). For phylogenetic analysis of each GST class, members of the sister class were used as an

[1]https://phytozome.jgi.doe.gov/pz/portal.html

[2]http://pfam.xfam.org/search

[3]http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

[4]http://www.drive5.com/muscle/

out-group. One-thousand bootstrap replicates were conducted to obtain confidence support.

Expression of GST Genes in *Capsella* Tissues

The expression patterns of *Capsella* GST members during growth under normal conditions were examined by reverse transcription PCR (RT-PCR). Seeds of *C. rubella* were germinated on agar plates (Murashige and Skoog, 1962) and vernalized at 4°C for 4 days. Then, the seeds were grown in growth chambers under normal conditions (14 h light/10 h dark cycle) at a temperature of 25°C/22°C (day/night). Seedling plants were transplanted to soil for 2 weeks and harvested for RT-PCR analysis. We isolated total RNA from rossette leaves, roots, and hypocotyl tissues of each plant and dry seeds using an Aurum Total RNA kit (Bio-Rad Laboratories). Total RNA was treated with RNase-free DNase I (Promega) and reverse transcribed into cDNA using a TaKaRa RNA PCR kit (AMV), version 3.0. Forty-nine specific primer pairs were designed (Supplementary Table S3). The *actin* gene (Carubv10013961m.g) was used as an internal control. PCR conditions were optimized to consist of an initial denaturation step of 3 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C and 30 s at 72°C, with a final extension of 5 min at 72°C. PCR products from each sample were analyzed on 1% agarose gel and were validated by DNA sequencing. Independent biological triplicates were used in all of the RT-PCR analyses.

Gene expression profiles of the *Capsella* GSTs were compared with expression data from *Arabidopsis* ecotype *Columbia-0* (*Col-0*; Schmid et al., 2005) available through the *Arabidopsis* eFP browser at BAR (Winter et al., 2007). The eFP browser was set to the developmental map, with absolute expression values for gene expression. In this study, genes with values below 20 units were considered to be not expressed (Winter et al., 2007). The microarray data sets used in this study include leaves at rosette stage (ATGE_89_A, ATGE_89_B and ATGE_89_C), roots at rosette stage (ATGE_9_A, ATGE_9_B and ATGE_9_C), hypocotyls at seedling stage (ATGE_2_A, ATGE_2_B and ATGE_2_C), and dry seeds (RIKEN-NAKABAYASHI1A and RIKEN-NAKABAYASHI1B).

Putative Promoter Sequence Analysis

Gene promoter sequences were extracted 1000 pb upstream of the transcriptional start site of each *Capsella* GST. Plant CARE database[5] was used to find putative *cis*–elements among the promoter sequences. Divergence between upstream sequences of each paralogous gene pairs was measured by the GATA program (Nix and Eisen, 2005), with window size set as seven and lower cutoff score 12 bit.

Expression and Purification of Recombinant *Capsella* GST Proteins

To investigate the enzymatic functions of *C. rubella* GST proteins, 24 tau, 11 phi, three DHAR, and three zeta GSTs were selected for protein expression analysis and purification. The primers used to construct the GST expression vectors are listed in Supplementary Table S4. The products were subcloned into pET-30a expression vectors (Novagen) to obtain a 6×His-tag at the N-terminus. The resulting plasmids, pET-30a/GSTs, were transformed into *E. coli* BL21 (DE3) and verified by sequencing. The transformed *E. coli* cells were cultured at 37°C and grown until the optical density (A_{600}) reached 0.5. A final concentration of 0.1 mM isopropyl-β-D-thiogalactopyranoside was added to each culture, and the cultures were incubated at 37°C or 20°C overnight. The cells were harvested by centrifugation (10,000 *g*, 3 min, 4°C), resuspended in binding buffer (20 mM sodium phosphate, 0.5 M NaCl, and 20 mM imidazole, pH 7.4), and disrupted by cold sonication. The resulting homogenate was subjected to centrifugation (10,000 × *g*, 10 min, 4°C) and the supernatant was loaded onto a Ni Sepharose High Performance column (GE Healthcare Bio-Sciences) that had been pre-equilibrated with binding buffer. The GST proteins that bound to the Ni Sepharose High Performance column were eluted with elution buffer (20 mM sodium phosphate, 0.5 M NaCl, and 0.5 M imidazole, pH 7.4). The particulate material, a small portion of the supernatant and the purified proteins were analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) consisting of a 10% separating gel and a 5% stacking gel.

Enzyme Assays of *Capsella* GST Proteins

The enzyme activity of plant GSTs was measured using the following six substrates: 1-chloro-2,4-dinitrobenzene (CDNB) and 4-nitrobenzyl chloride (NBC), as described by Habig et al. (1974); 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (NBD-Cl), as described by Ricci et al. (1994); and cumene hydroperoxide (Cum-OOH), dehydroascorbate (DHA), and diphenyl ethers (Fluorodifen), as described by Edwards and Dixon (2005). All assays were carried out at 25°C. Protein concentrations were determined by measuring the absorbance at 280 nm.

RESULTS

Identification of the GST Genes from the *C. rubella* Genome

Forty-nine full-length genes encoding putative cytosolic GST proteins were identified in the *C. rubella* genome (Supplementary Table S1). Among these 49 genes, two genes (*CrGSTF3* and *CrGSTU9*) were considered to be putative pseudogenes because one contained a frame shift disrupting the coding region and the other contained a premature stop codon. After revising the frame shifts by deleting two nucleotides or removing the stop codon, these two full-length sequences were included in the phylogenetic and gene expression analyses. Based on a conserved domain analysis, these 49 GST candidates were divided into eight classes. The tau and phi GSTs were most numerous, with 25 and 12 copies, respectively. The DHAR and zeta classes each contained three members. Both the lambda and EF1Bγ classes were each represented by two members, and the theta and TCHQD classes only had one member each.

[5]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

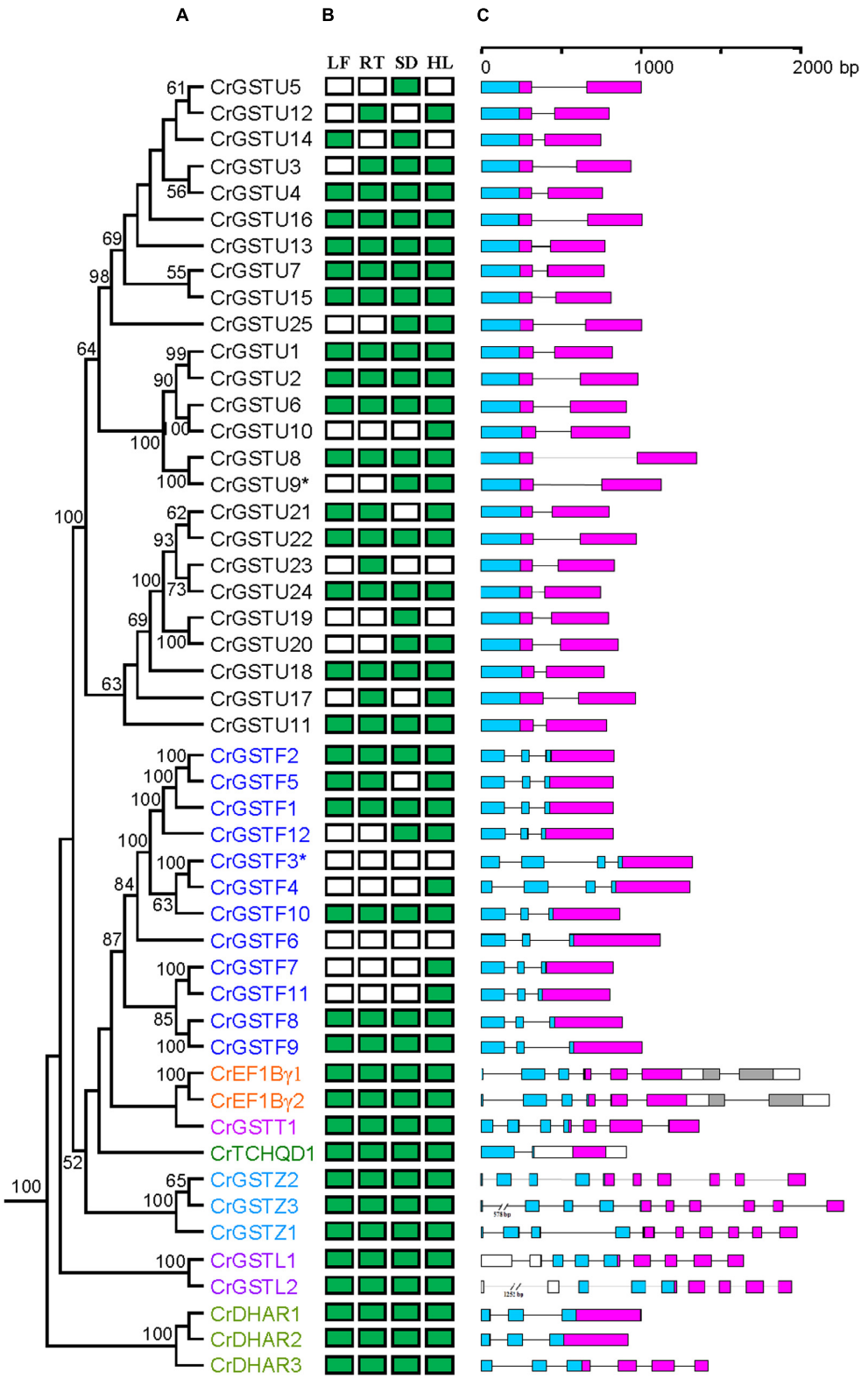


FIGURE 1 | Phylogenetic relationships among *Capsella* GSTs (A), their expression patterns (B) and gene structures (C). Numbers at each node in the phylogenetic tree represent bootstrap values, and only values higher than 50% are shown. The Glutathione *S*-transferase (GST) genes belonging to different classes are indicated with different colors. In **(B)**, the green box indicates positive detection of gene expression in leaf (LF), root (RT), seed (SD), and hypocotyls (HL) under normal growth conditions. In **(C)**, the GST N- and C-terminal domain are highlighted by blue and purple boxes, respectively. Introns are shown as lines

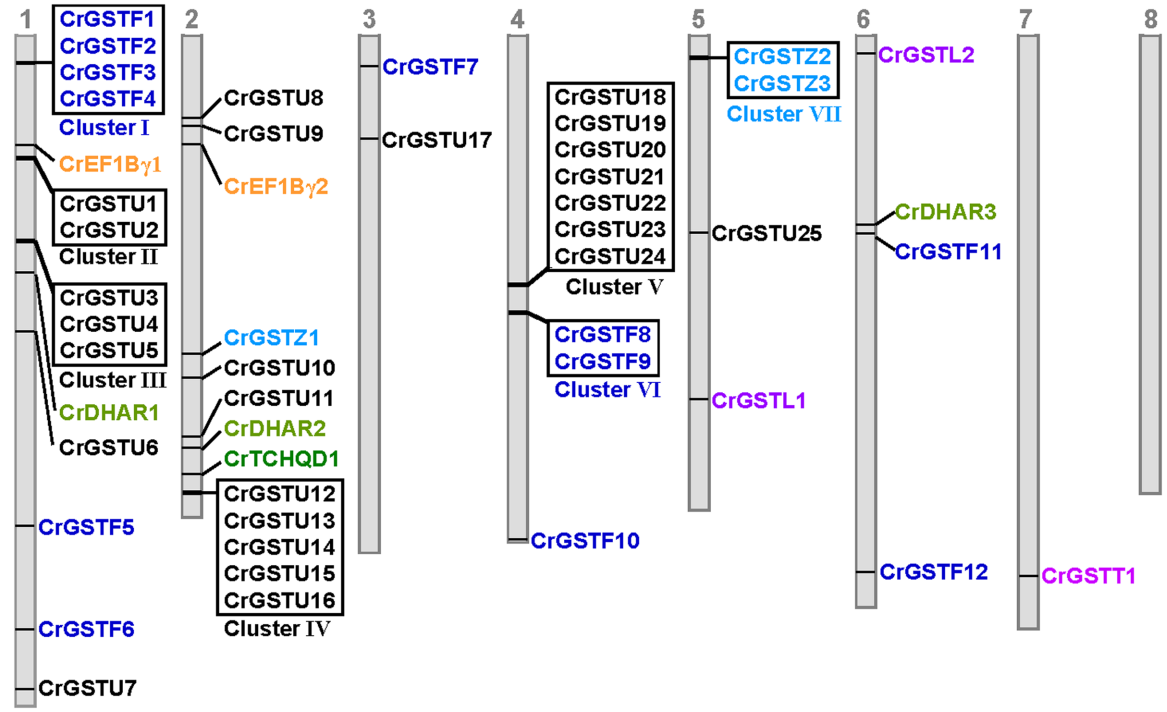


FIGURE 2 | Genomic localization of *Capsella* GSTs. The numbered gray bars represent the eight *Capsella* chromosomes. The seven tandem-arrayed GST clusters are shown in boxes

Conserved gene structures were identified within each GST class. All 25 tau GST genes contained a one-intron/two-exon structure (**Figure 1C**). The intron positions were highly conserved among tau GST genes, and their lengths ranged from 70 to 642 bp. Ten of 12 phi GSTs had a two-intron/three-exon structure with a highly conserved first intron position. *CrGSTF3* and *CrGSTF4* contained a three-intron/four-exon structure. For the EF1Bγ, lambda, and zeta classes, the exon-intron architectures were conserved within each class, with each including 7, 9, and 10 exons, respectively.

Genomic Organization of the *Capsella* GST Gene Family

The genomic locations of 49 full-length GSTs were assigned to all of the *Capsella* chromosomes except for chromosome 8 (**Figure 2**). The distribution of the GST genes among the chromosomes was obviously heterogeneous. Seven clusters (clusters I, II, III, IV, V, VI, and VII) with relatively high densities of GSTs were discovered on four chromosomes. In total, 51% of *Capsella* GST genes were organized in tandem repeats, indicating that tandem duplications significantly contributed to the expansion of the *Capsella* GST gene family.

Among the seven clusters, cluster V was the largest, consisting of seven tau GSTs in a 25-kb region on chromosome 4. These seven GST genes were oriented in the same direction. Cluster IV contained five tau GST genes that were clustered in a 10-kb region on chromosome 2. Three clusters (cluster I, II, and III) were located on chromosome 1. Cluster I harbored four phi GSTs arranged head-to-head in tandem in an 8-kb region. Clusters II and III contained two and three tau GSTs, respectively. Cluster VI contained two phi GST genes, and cluster VII contained two zeta GSTs.

Phylogenetic Relationship of *Capsella* and *Arabidopsis* GST Gene Families

To investigate the lineage-specific expansion of GST genes in the *Capsella* genome, we performed a phylogenetic analysis of all GSTs from *Capsella* and *Arabidopsis*. *Capsella* and *Arabidopsis* had 25 and 28 tau GST genes, respectively. There were at least 27 ancestral GST genes in the most recent common ancestor (MRCA) of *Capsella* and *Arabidopsis* (**Figures 3** and **4**). After the split, both *Capsella* and *Arabidopsis* gained one gene. However, *Capsella* had lost three genes, resulting in fewer tau GST genes in *Capsella* compared with *Arabidopsis*.

Capsella and *Arabidopsis* had 12 and 13 phi GST genes, respectively. There were at least 12 ancestral GST genes in the MRCA of *Capsella* and *Arabidopsis* (**Figures 3** and **5**). After the split, *Capsella* and *Arabidopsis* gained one and two genes, respectively. Additionally, *Capsella* and *Arabidopsis* each lost one gene. Thus, *Capsella* contains one less phi GST gene than *Arabidopsis*.

For DHAR and lambda GSTs, there were at least four and three ancestral GST genes, respectively, in the MRCA of *Capsella* and *Arabidopsis* (**Figures 3** and **5**). After the split, *Capsella* lost a DHAR and lambda GST; however, the *Arabidopsis* genome did not lose any DHAR or lambda GST genes.

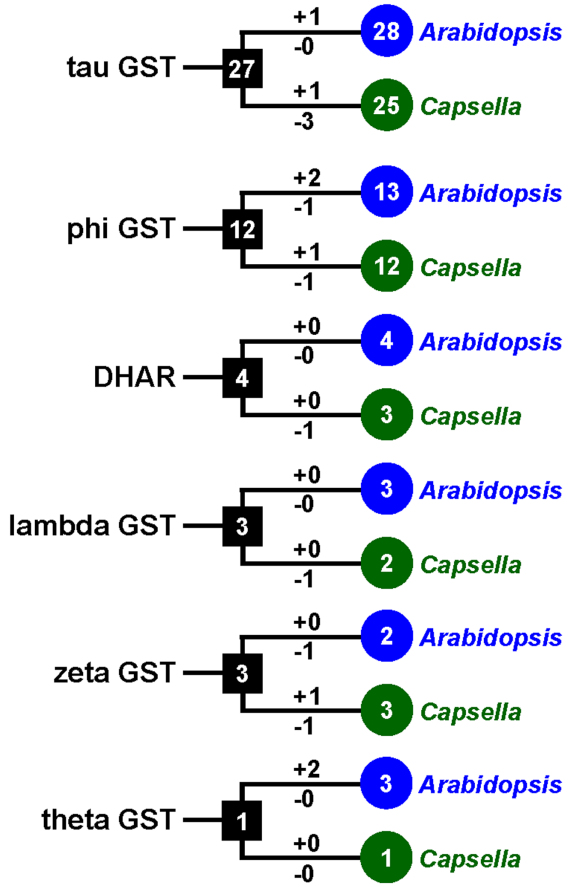


FIGURE 3 | The copy number changes of *Capsella* and *Arabidopsis* GSTs. Numbers in circles and rectangles represent the numbers of GST genes in extant and ancestral species, respectively. Numbers on branches with plus and minus symbols represent the numbers of gene gains and losses, respectively.

For zeta GST, the MRCA of *Capsella* and *Arabidopsis* had at least three ancestral zeta GSTs (**Figures 3** and **5**). After the split, the *Capsella* genome gained and lost one gene, and thus, *Capsella* still contains three zeta GSTs. *Arabidopsis* did not gain new copies. On the contrary, one zeta GST gene was lost, thus, the *Arabidopsis* genome contains two zeta GSTs.

For theta class GSTs, at least one ancestral GST gene was noted in the MRCA of the two species (**Figure 3**). Two duplication events in theta class GSTs were found in the *Arabidopsis* lineage after the split from *Capsella* (**Figure 5**). For TCHQD and EF1B class GSTs, we did not identify any gene gain or loss events after the split of these two species (**Figure 5**).

Expression Patterns of the *Capsella* GST Gene Family

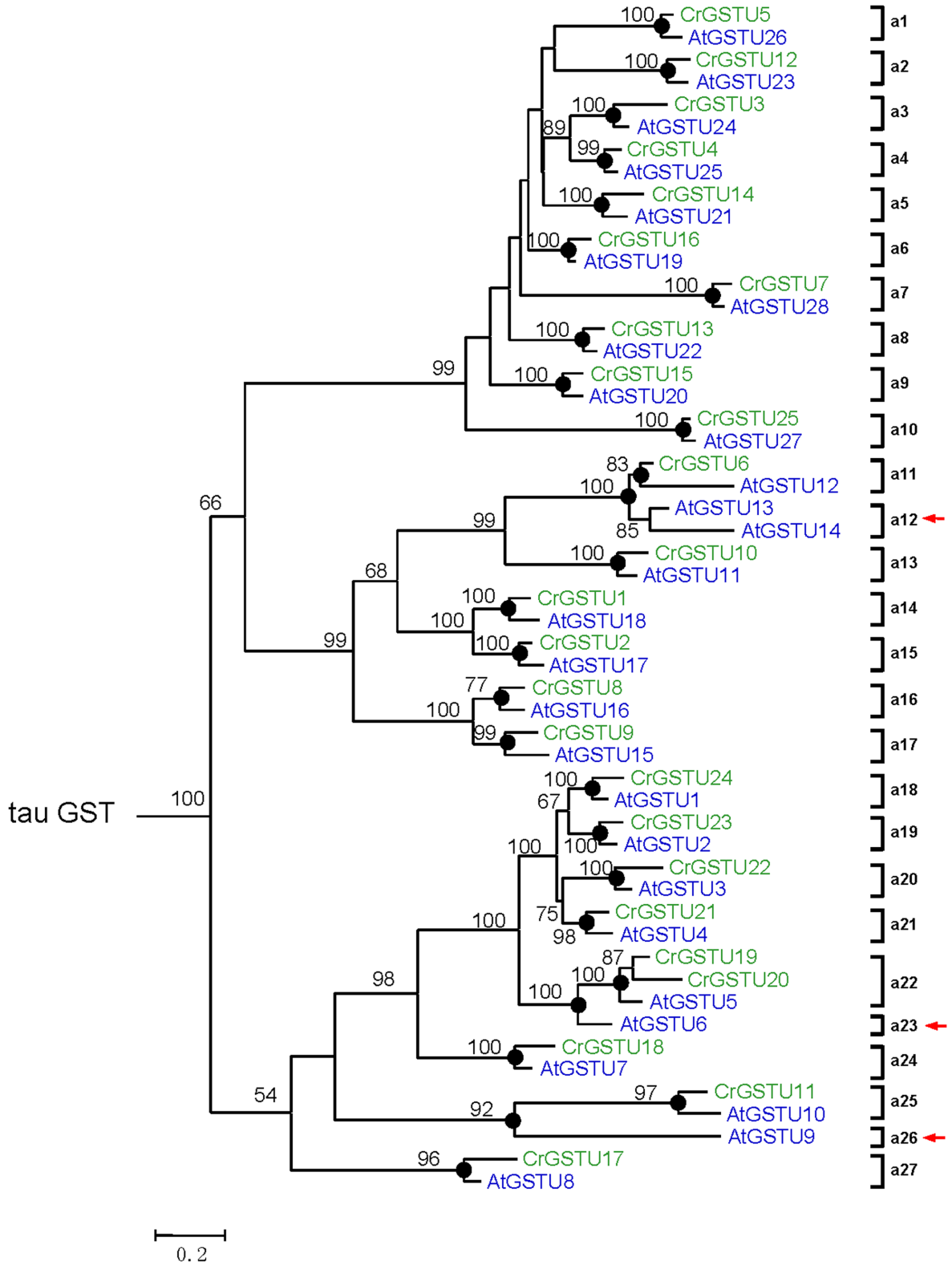
We examined the tissue-specific expression patterns of all 49 *Capsella* GSTs in four tissues, including leaves, roots, seeds, and hypocotyl zones using RT-PCR (**Figure 1B**). The expression patterns of the six minor class GSTs (*CrEF1Bγ1* and *2*, *CrGSTT1*, *CrTCHQD1*, *CrGSTZ1*, *2* and *3*, *CrGSTL1* and *2*, and *CrDHAR1*, *2* and *3*) were homogenous, as these GSTs were expressed in all tissues examined in this study. Expression divergences were observed among the tau and phi class GSTs (**Figure 1B**). Of the 25 tau GSTs, 13 (*CrGSTU1*, *2*, *4*, *6*, *7*, *8*, *11*, *13*, *15*, *16*, *18*, *22*, and *24*) were expressed in all tissues examined, and 12 genes (*CrGSTU3*, *5*, *9*, *10*, *12*, *14*, *17*, *19*, *20*, *21*, *23*, and *25*) were selectively expressed. *CrGSTU23* was exclusively expressed in root tissues. *CrGSTU5* and *CrGSTU19* were exclusively expressed in seed tissues, whereas *CrGSTU10* was only noted in hypocotyl zones. For the 12 phi GSTs, five (*CrGSTF1*, *2*, *8*, *9*, and *10*) were expressed in all the tissues examined, and five (*CrGSTF4*, *5*, *7*, *11*, and *12*) were selectively expressed. *CrGSTF3* and *CrGSTF6* were not expressed in any of the tissues examined (**Figure 1B**), which suggests that these two genes might exhibit loss of gene function by pseudogenization.

The expression profiles of 23 tau and 8 phi orthologous GSTs in *Capsella* and *Arabidopsis* revealed a high degree of divergence (**Figure 6**). In total, 7 of the 31 orthologs displayed similar expression patterns, whereas the remaining 24 orthologs exhibited considerable expression divergence in some tissues. For example, *AtGSTU21* was not detected in any tissue of *Arabidopsis* (Goda et al., 2008; Kram et al., 2009), but its orthologous gene *CrGSTU14* was expressed in leaf and seed tissues (**Figures 1B** and **6**).

Potential regulatory motifs analysis using PlantCARE (Plant *cis*-acting regulatory element database) revealed a number of *cis*-elements in the promoter sequences of *Capsella* GST genes (Supplementary Table S6). These motifs were divided into at least eight functional categories, such as core promoters, ABA/abiotic stress, light, phytohormones, pathogen/elicitor/wound responsive elements as well as elements responsible for metabolism regulation, developmental stage, and organ specific expressions. The result showed considerable differences in the regulatory elements among the *Capsella* GSTs and within the subfamilies. Comparative analysis of upstream regions of close paralogs showed divergence, although there are conserved regions (Supplementary Figure S1), indicating that rapid divergence has occurred in the regulatory regions. Further experimental validation step is required to assess the changes in the *cis*-elements that are responsible for the expression diversity in GST genes.

Substrate Specificities and Activities of *Capsella* GSTs]

In the *Capsella* genome, the tau and phi GSTs are most numerous, with 25 and 12 copies, respectively. The DHAR and zeta classes each contain three members. Thus, in this study, we selected tau, phi, DHAR, and zeta GSTs to express and purify GST proteins. Except for two pseudogenes (CrGSTU9 and CrGSTF3), Forty-one genes were cloned into expression vector pET-30a. Twenty-five of the 41 recombinant proteins were expressed as soluble proteins in *E. coli*, whereas the other 16 were insoluble. To determine the enzyme activity and substrate specificity of the soluble proteins, six substrates were selected: CDNB, NBD-Cl, NBC, Cum-OOH, and DHA.


tau GST
100
66
99
68
99
54
98
100
100
100

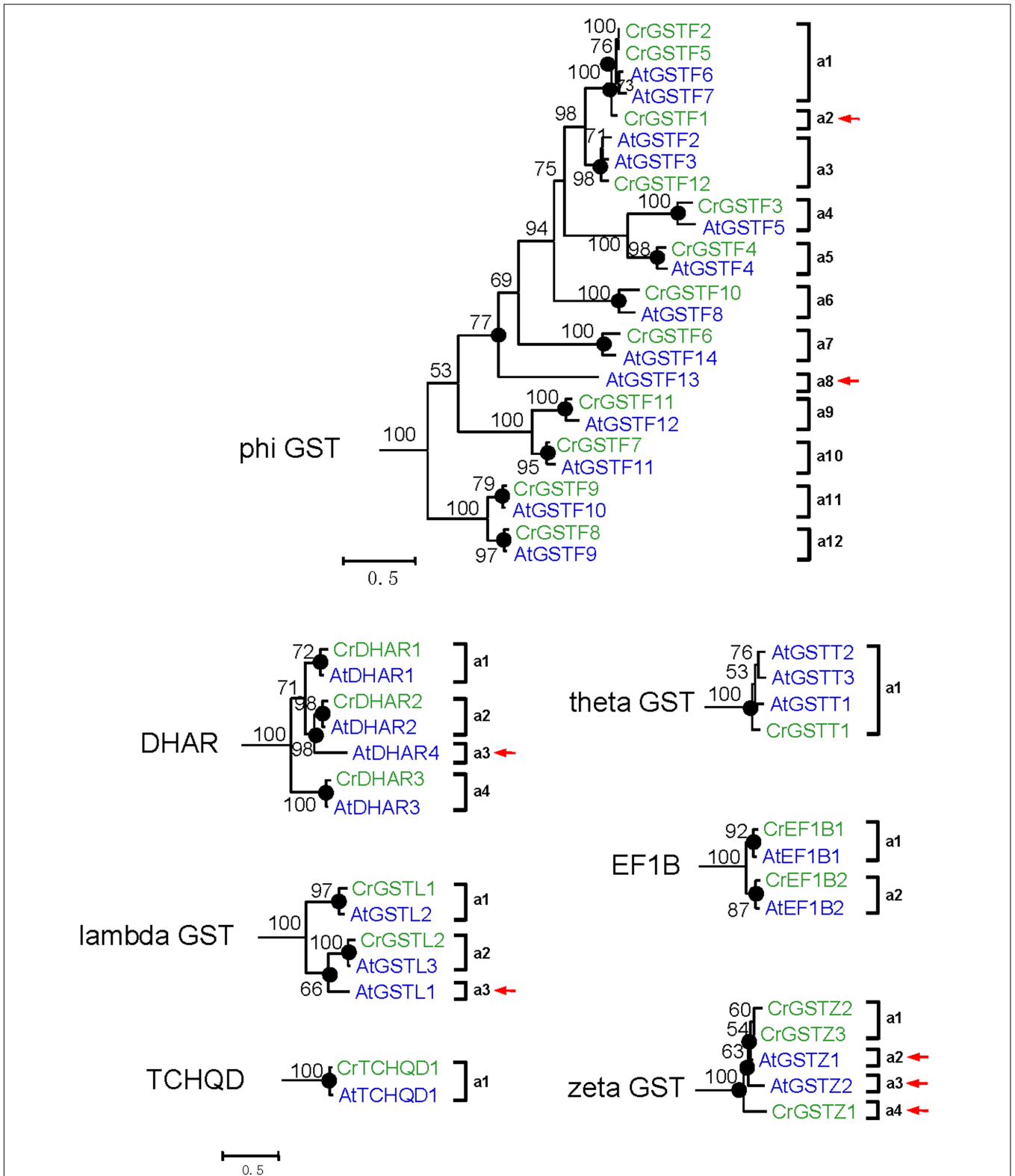


FIGURE 5 | Phylogenetic relationships of the *Capsella* and *Arabidopsis* phi and all six minor class GSTs. Numbers at each node in the phylogenetic tree represent bootstrap values, and only values higher than 50% are shown. *Capsella* and *Arabidopsis* GSTs are indicated in green and blue, respectively. The nodes that represent the most recent common ancestral genes before the *Capsella* and *Arabidopsis* split are indicated by black circles. Clades that contain only *Capsella* or *Arabidopsis* GSTs are indicated by red arrows

For the tau GSTs, all 14 proteins showed specific activity toward CDNB, 11 toward NBD-Cl, nine toward Cum-OOH, and seven toward NBC and fluorodifen (**Figure 7**). Among the 14 tau GSTs, two (CrGSTU2 and 4) had enzymatic activity toward all five of the substrates. Five proteins exhibited activity toward four substrates, and four toward three substrates. Among the tau GSTs, CrGSTU4 showed the highest enzymatic activity toward CDNB, CrGSTU16 toward NBD-Cl, CrGSTU7 toward NBC, CrGSTU21 toward Cum-OOH, and CrGSTU2 toward fluorodifen. For the 12 phi GSTs, only five proteins were expressed as soluble proteins in *E. coli*. Among these five proteins, CrGSTF10 exhibited boarder substrate spectra and enzymatic activities for four substrates. CrGSTF1 did not exhibit any activity toward any of the tested substrates. All three DHAR GSTs exhibited activity toward DHA. Compared with CrDHAR1 and 3, CrDHAR2 had noticeably reduced enzymatic activity toward DHA, whereas CrDHAR2 exhibited activity toward Cum-OOH.

Substantial variations in specific activities toward different substrates were noted among the members of tandem-arrayed GST clusters. For example, CrGSTU12, 15, and 16 belong to cluster IV. CrGSTU16 displayed a much broader substrate spectrum than did CrGSTU12 and 15. Although CrGSTU12 and 15 shared a similar substrate spectrum, their specific activities toward NBD-Cl varied 10-fold (**Figure 7**).

The enzymatic activities of orthologous GSTs in *Capsella* and *Arabidopsis* also displayed variations. We made a comparison of enzyme specificity toward CDNB and Cum-OOH between orthologous GSTs (Dixon et al., 2009). For example, CrGSTU7, CrGSTU12, CrGSTU15, and CrGSTU18 had enzymatic activity for CDNB but no activity for Cum-OOH, whereas their orthologs in *Arabidopsis* had enzymatic activity for both substrates (**Figures 6** and **7**).

DISCUSSION

Functional divergence of duplicated genes is a major factor promoting their retention in the genome (Ohno, 1970; Zhang, 2003). Many theoretical models have been proposed to explain the mechanisms leading to the divergence include sub-functionalization, neo-functionalization, and dosage-balance model, etc (Ohno, 1970; Hughes, 1994; Force et al., 1999; Walsh, 2003; Moore and Purugganan, 2005; Veitia et al., 2008; Innan and Kondrashov, 2010). However, our understanding of the mechanisms driving the evolution of a large and functionally heterogeneous gene family is limited. Because to determine whether the duplicates have identical, similar, or different functions requires comprehensive examination of the functions of each gene product, while this approach is useful at a single gene level, genome-scale analyses of functional divergence of a supergene family are unfeasible. Our study combined bioinformatics and experimental approaches to explore the functional diversification of GST gene family at different levels of genomic organization: among subfamily classes, within tandem clusters, in paralogous and orthologous gene pairs.

Genome annotation identified 49 full-length GST genes from the *C. rubella* genome, which were divided into eight classes. Extensive study has showed that tau and phi classes were the most abundant with wide interspecific variation in copy number in plants (Lan et al., 2009, 2013; Dixon and Edwards, 2010a; Jain et al., 2010; Liu et al., 2013, 2015). For instance, our study and previous studies showed that tau GSTs was not found in moss and green algae, whereas it's ubiquitous in tracheophytes (25–62 copies). Seventeen phi GSTs were found in rice while only two were represented in *S. moellendorffii*. However, other six classes remain comparatively small, with only 1–5 members. Comparison of copy numbers among the classes indicated that they might follow distinct evolutionary paths. Why did extensive expansion of tau and phi classes occur? A possible explanation is functional requirement. Tau and phi GSTs play an important role in the detoxification of xenobiotics and defense responses against both biotic and abiotic stresses (Loyall et al., 2000; Karavangeli et al., 2005; Benekos et al., 2010; Dixon et al., 2011; Jha et al., 2011; Cummins et al., 2013). Thus, the large scale expansion within tau and phi classes might provide defense against a broader range of xenobiotics and facilitated their tolerance to various environmental hazards. Our study exhibited extensive diversification in enzyme substrate specificity and transcript expression in tau and phi classes. This might further support the diversification in response to a set of changing substrates and regulatory properties.

The rapid expansion of GST gene family in plants is largely the result of the expansion of tau and phi classes. In the *C. rubella* genome, 17 of the 25 (68%) tau GSTs consisted of four clusters, and six of the 12 (50%) phi GSTs consisted of two clusters, indicating that tandem duplication considerably contributed to the expansion of tau and phi GSTs in the *C. rubella* genome. Previous studies also indicate that tandem duplication played important roles in the expansion of tau and phi GSTs in poplar, soybean, *Arabidopsis,* and rice genomes (Dixon et al., 2002b; Soranzo et al., 2004; Lan et al., 2009; Liu et al., 2015). Why have so many duplicate genes been retained for such a long time in the *C. rubella* genome? To investigate this question, we examined the seven tandem-arrayed clusters (Cluster I–VII). We found two categories of expression patterns. In the first, all the members in each cluster were expressed in all tissues. This pattern was observed in tau cluster II (*CrGSTU1/2*), phi cluster VI (*CrGSTF8/9*), and zeta cluster VII (*CrGSTZ2/3*; **Figure 1**). In the second category, found in tau cluster III (*CrGSTU3-5*), IV (*CrGSTU12-16*), V (*CrGSTU18-24*), and phi cluster I (*CrGSTF1-4*), some copies were expressed in all tissues, some had restricted tissue-specific expression or were not expressed in any tissue examined (**Figure 1B**). When enzyme assays were examined, no GST proteins in clusters showed identical enzymatic activities and specificities toward different substrates (**Figure 7**). Through this integrated approach, we found that rapid divergence has occurred in the regulatory regions of genes and in their biochemical properties within clusters, suggesting that partial sub-functionalization has indeed taken place. This seems to be an important factor promoting the duplicated GSTs' retention in the genome.

A major challenge in comparative genomics is to find sufficient functional differences between species. However, it remains challenging in *Arabidopsis* and other plants, partly due to

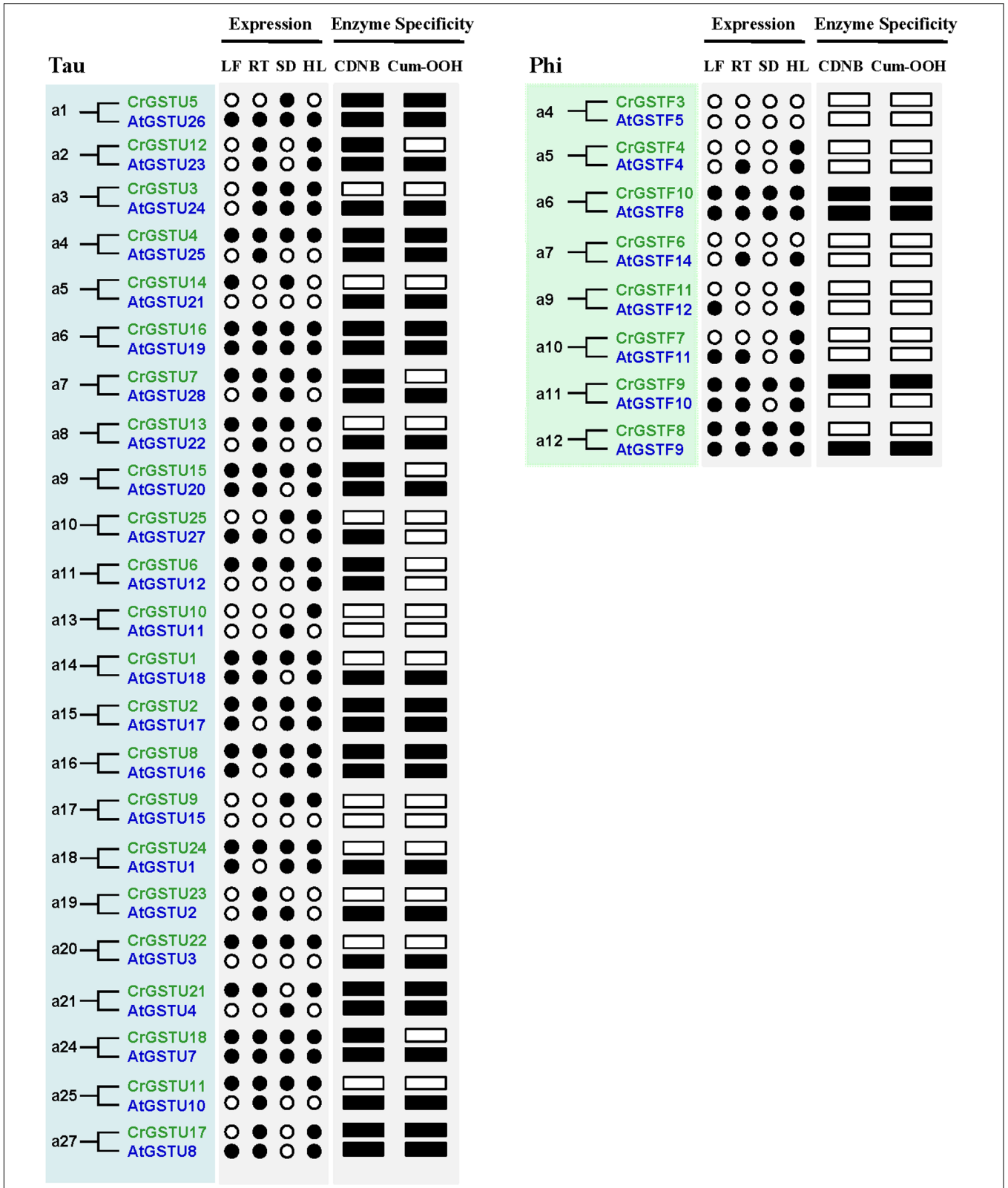
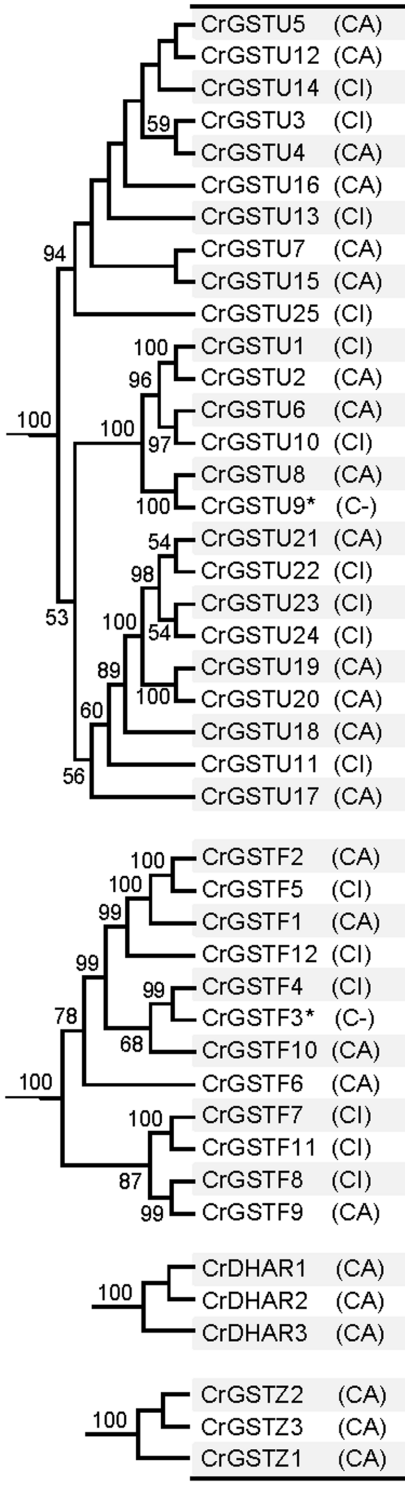


FIGURE 6 | Expression and functional divergence between ortholog gene pairs in *Capsella* and *Arabidopsis*. The black circle and box indicate positive detection of gene expression in the corresponding tissue and specific activity toward 1-chloro-2,4-dinitrobenzene (CDNB) or cumene hydroperoxide (Cum-OOH), respectively.



GSTs		Specific Activity (μmol/min per mg) to Each Substrate					
		CDNB	NBD-Cl	NBC	Cum-OOH	Fluorodifen	DHA
CrGSTU5	(CA)	0.91±0.03	3.36±0.10	n.d.	0.06±0.01	n.d.	n.d.
CrGSTU12	(CA)	1.55±0.05	0.18±0.01	n.d.	n.d.	n.d.	n.d.
CrGSTU14	(CI)						
CrGSTU3	(CI)						
CrGSTU4	(CA)	17.07±0.16	7.92±0.27	0.59±0.02	0.17±0.02	0.44±0.02	n.d.
CrGSTU16	(CA)	6.60±0.11	44.85±2.43	0.63±0.11	0.03±0.01	n.d.	n.d.
CrGSTU13	(CI)						
CrGSTU7	(CA)	0.85±0.04	n.d.	1.80±0.10	n.d.	13.33±0.04	n.d.
CrGSTU15	(CA)	1.69±0.05	1.84±0.03	n.d.	n.d.	n.d.	n.d.
CrGSTU25	(CI)						
CrGSTU1	(CI)						
CrGSTU2	(CA)	4.39±0.02	5.55±0.06	0.70±0.03	0.35±0.02	21.12±1.00	n.d.
CrGSTU6	(CA)	0.12±0.01	0.03±0.00	n.d.	n.d.	n.d.	n.d.
CrGSTU10	(CI)						
CrGSTU8	(CA)	1.12±0.02	20.14±0.48	n.d.	0.12±0.01	3.84±0.18	n.d.
CrGSTU9*	(C-)						
CrGSTU21	(CA)	0.82±0.04	1.03±0.05	1.08±0.03	0.46±0.04	n.d.	n.d.
CrGSTU22	(CI)						
CrGSTU23	(CI)						
CrGSTU24	(CI)						
CrGSTU19	(CA)	3.92±0.59	n.d.	1.08±0.16	0.20±0.02	10.94±0.26	n.d.
CrGSTU20	(CA)	1.20±0.04	n.d.	n.d.	0.05±0.00	1.54±0.09	n.d.
CrGSTU18	(CA)	0.03±0.01	0.03±0.03	n.d.	n.d.	0.11±0.02	n.d.
CrGSTU11	(CI)						
CrGSTU17	(CA)	0.82±0.01	0.27±0.00	1.29±0.01	0.17±0.01	n.d.	n.d.
CrGSTF2	(CA)	0.08±0.00	n.d.	n.d.	0.68±0.03	n.d.	n.d.
CrGSTF5	(CI)						
CrGSTF1	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
CrGSTF12	(CI)						
CrGSTF4	(CI)						
CrGSTF3*	(C-)						
CrGSTF10	(CA)	5.65±0.17	0.56±0.02	1.13±0.02	0.16±0.02	n.d.	n.d.
CrGSTF6	(CA)	n.d.	n.d.	1.96±0.07	n.d.	n.d.	n.d.
CrGSTF7	(CI)						
CrGSTF11	(CI)						
CrGSTF8	(CI)						
CrGSTF9	(CA)	1.18±0.06	0.03±0.00	n.d.	0.14±0.01	n.d.	n.d.
CrDHAR1	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	39.78±1.43
CrDHAR2	(CA)	n.d.	n.d.	n.d.	0.09±0.01	n.d.	0.73±0.01
CrDHAR3	(CA)	n.d.	n.d.	n.d.	n.d.	n.d.	30.17±0.50
CrGSTZ2	(CA)	n.d.	0.06±0.01	n.d.	0.03±0.00	n.d.	n.d.
CrGSTZ3	(CA)	n.d.	n.d.	n.d.	0.02±0.01	n.d.	n.d.
CrGSTZ1	(CA)	n.d.	n.d.	n.d.	0.05±0.01	n.d.	n.d.

FIGURE 7 | Specific activities of the *Capsella* tau, phi, dehydroascorbate reductase (DHAR), and zeta class GSTs toward six substrates (Mean ± SD obtained from at least three independent determinations). C, successfully cloned; A, purified GST protein assayed; I, recombinant protein totally insoluble; dash, analysis not performed; n.d., no activity detected.

technical limits and potential functional redundancy within the family (Sappl et al., 2009). We identified 23 tau and 8 phi orthologous GSTs in the two relatives, and most of the gene pairs exhibited variations at expression and biochemical level (**Figure 6**), indicating that their functions may have evolved after the split. For example, *AtGSTF12* showed high expression in senescent leaf and was demonstrated to be involved in flavonoid metabolism (Kitamura et al., 2004; Dixon and Edwards, 2010a). But its orthologous gene *CrGSTF11* was not detected in leaf tissue (**Figure 1B**). AtGSTU25 and AtGSTU28 have the highest activity in tau class when assayed with CDNB or Cum-OOH as substrates (Dixon et al., 2009), whereas their orthologs CrGSTU4 and CrGSTU7 have low activity for Cum-OOH and CDNB, respectively (**Figure 7**). *AtGSTU20* was showed to interact with *Far-red (FR) insensitive 219* (*FIN219*) in response to light and could regulate cell elongation and plant development (Chen et al., 2007). We detected some light responsive elements in the promoter of *CrGSTU15* (Supplementary Table S6), suggesting that *CrGSTU15* may also involve in light regulation. In addition, 5 of the 31 orthologs displayed similar patterns (**Figure 6**). *AtGSTU19* and *CrGSTU16* displayed similar expression and substrate spectrum. *AtGSTU19* showed tolerance to salt, drought, and methyl viologen stresses (Xu et al., 2015). Several *cis*-acting elements involved in abscisic acid, anaerobic, heat, low-temperature, drought, defense, and phytohormones responsiveness were indentified in *CrGSTU16* as well, suggesting that this gene may also be induced by several stimuli. However, *AtGSTF8* and *CrGSTF10* exhibited a different example: As an enzyme, *AtGSTF8* was the most active member in phi class when assayed with CDNB and Cum-OOH (Dixon et al., 2009). Its expression was strongly induced by salicylic acid and H_2O_2 in root tissue, and *ocs* element in the promoter region played an important role (Chen and Singh, 1999). Unlike *AtGSTF8*, *CrGSTF10* didn't contain *ocs*-like element and its specific activity toward Cum-OOH was low (Supplementary Table S6 and **Figure 7**). Protein subcellular relocalization is also considered as another form of functional divergence (Marques et al., 2008; Qian and Zhang, 2009; Wang et al., 2009). *AtGSTU12* is the only tau class GST localized entirely to the nucleus, containing a putative nuclear localization signal KKRKK (Takahashi et al., 1995). But we did not find this signal in *CrGSTU6*. All these results suggested that functional divergence previously had occurred in the two lineages after the split.

CONCLUSION

In this study, we characterized the complete set of GST gene family in *C. rubella* genome. By phylogenetic and functional analysis, we compared it to that in *Arabidopsis*. We examined the gene gain and loss events after the divergence of the relatives. Also, we evaluated the functional divergences of recently expanded GSTs and orthologs. Through these analyses, we were able to draw a picture illustrating how gene duplication and sub-functionalization influence the divergence, retention, and functions of GST genes in the *Capsella* genome. Furthermore, by extending genome-wide comparison analysis of GST gene family with more species in the Brassicaceae, the study will provide a comprehensive overview of the evolutionary history of a large gene family among lineages and mechanism of functional diversification and retention of duplicates.

AUTHOR CONTRIBUTIONS

TL and QZ designed the study. GH, CG, and QC performed the experiments. TL, XG, and WL analyzed the data. TL wrote the paper.

ACKNOWLEDGMENTS

We thank Ya-Long Guo (Institute of Botany, Chinese Academy of Sciences) for the seeds of *Capsella rubella*. This work was supported by the grant from the National Natural Science Foundation of China (No. 31200171) and the National Science Foundation for Distinguished Young Scholars of China (No. 31425006).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01325

REFERENCES

Acarkan, A., Rossberg, M., Koch, M., and Schmidt, R. (2000). Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. *Plant J.* 23, 55–62. doi: 10.1046/j.1365-313x.2000.00790.x

Basantani, M., and Srivastava, A. (2007). Plant glutathione transferases - a decade falls short. *Can J. Bot.* 85, 443–456. doi: 10.1139/B07-033

Benekos, K., Kissoudis, C., Nianiou-Obeidat, I., Labrou, N., Madesis, P., Kalamaki, M., et al. (2010). Overexpression of a specific soybean GmGSTU4 isoenzyme improves diphenyl ether and chloroacetanilide herbicide tolerance of transgenic tobacco plants. *J. Biotechnol.* 150, 195–201. doi: 10.1016/j.jbiotec.2010.07.011

Boivin, K., Acarkan, A., Mbulu, R. S., Clarenz, O., and Schmidt, R. (2004). The *Arabidopsis* genome sequence as a tool for genome analysis in Brassicaceae. *A* comparison of the *Arabidopsis* and *Capsella rubella* genomes. *Plant Physiol.* 135, 735–744. doi: 10.1104/pp.104.040030

Chen, I. C., Huang, I. C., Liu, M. J., Wang, Z. G., Chung, S. S., and Hsieh, H. L. (2007). Glutathione S-transferase interacting with far-red insensitive 219 is involved in phytochrome A-mediated signaling in *Arabidopsis*. *Plant Physiol.* 143, 1189–1202. doi: 10.1104/pp.106.094185

Chen, W., and Singh, K. B. (1999). The auxin, hydrogen peroxide and salicylic acid induced expression of the *Arabidopsis* GST6 promoter is mediated in part by an ocs element. *Plant J.* 19, 667–677. doi: 10.1046/j.1365-313x.1999.00560.x

Chen, Z., and Gallie, D. R. (2006). Dehydroascorbate reductase affects leaf growth, development, and function. *Plant Physiol.* 142, 775–787. doi: 10.1104/pp.106.085506

Cummins, I., Cole, D. J., and Edwards, R. (1999). A role for glutathione transferases functioning as glutathione peroxidases in resistance to multiple herbicides in black-grass. *Plant J.* 18, 285–292. doi: 10.1046/j.1365-313X.1999.00452.x

Cummins, I., Dixon, D. P., Freitag-Pohl, S., Skipsey, M., and Edwards, R. (2011). Multiple roles for plant glutathione transferases in xenobiotic detoxification. *Drug Metab. Rev.* 43, 266–280. doi: 10.3109/03602532.2011.552910

Cummins, I., Wortley, D. J., Sabbadin, F., He, Z., Coxon, C. R., Straker, H. E., et al. (2013). Key role for a glutathione transferase in multiple-herbicide resistance in grass weeds. *Proc. Natl. Acad. Sci. U.S.A.* 110, 5812–5817. doi: 10.1073/pnas.1221179110

Dixon, D. P., Davis, B. G., and Edwards, R. (2002a). Functional divergence in the glutathione transferase superfamily in plants. Identification of two classes with putative functions in redox homeostasis in *Arabidopsis thaliana*. *J. Biol. Chem.* 277, 30859–30869.

Dixon, D. P., and Edwards, R. (2010a). Glutathione transferases. *Arabidopsis Book* 8:e0131. doi: 10.1199/tab.0131

Dixon, D. P., and Edwards, R. (2010b). Roles for stress-inducible lambda glutathione transferases in flavonoid metabolism in plants as identified by ligand fishing. *J. Biol. Chem.* 285, 36322–36329. doi: 10.1074/jbc.M110.164806

Dixon, D. P., Lapthorn, A., and Edwards, R. (2002b). Plant glutathione transferases. *Genome Biol.* 3, 1–10. doi: 10.1186/gb-2002-3-3-reviews3004

Dixon, D. P., Hawkins, T., Hussey, P. J., and Edwards, R. (2009). Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. *J. Exp. Bot.* 60, 1207–1218. doi: 10.1093/jxb/ern365

Dixon, D. P., Skipsey, M., and Edwards, R. (2010). Roles for glutathione transferases in plant secondary metabolism. *Phytochemistry* 71, 338–350. doi: 10.1016/j.phytochem.2009.12.012

Dixon, D. P., Steel, P. G., and Edwards, R. (2011). Roles for glutathione transferases in antioxidant recycling. *Plant Signal. Behav.* 6, 1223–1227. doi: 10.4161/psb.6.8.16253

Edwards, R., and Dixon, D. P. (2005). Plant glutathione transferases. *Methods Enzymol.* 401, 169–186. doi: 10.1016/S0076-6879(05)01011-6

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., and Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531–1545.

Frova, C. (2003). The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. *Physiol. Plant.* 119, 469–479. doi: 10.1046/j.1399-3054.2003.00183.x

Frova, C. (2006). Glutathione transferases in the genomics era: new insights and perspectives. *Biomol. Eng.* 23, 149–169. doi: 10.1016/j.bioeng.2006.05.020

Goda, H., Sasaki, E., Akiyama, K., Maruyama-Nakashita, A., Nakabayashi, K., Li, W., et al. (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J.* 55, 526–542. doi: 10.1111/j.0960-7412.2008.03510.x

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52, 696–704. doi: 10.1080/10635150390235520

Guo, Y. L., Bechsgaard, J. S., Slotte, T., Neuffer, B., Lascoux, M., Weigel, D., et al. (2009). Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5246–5251. doi: 10.1073/pnas.0808012106

Guo, Y. L., Todesco, M., Hagmann, J., Das, S., and Weigel, D. (2012). Independent FLC Mutations as Causes of Flowering-Time Variation in *Arabidopsis thaliana* and *Capsella rubella*. *Genetics* 192, 729–739. doi: 10.1534/genetics.112.143958

Habig, W. H., Pabst, M. J., and Jakoby, W. B. (1974). Glutathione S-transferases. The first enzymatic step in mercapturic acid formation. *J. Biol. Chem.* 249, 7130–7139.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* 41, 95–98.

Hayes, J. D., Flanagan, J. U., and Jowsey, I. R. (2005). Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* 45, 51–88. doi: 10.1146/annurev.pharmtox.45.120403.095857

Holm, P. J., Bhakat, P., Jegerschold, C., Gyobu, N., Mitsuoka, K., Fujiyoshi, Y., et al. (2006). Structural basis for detoxification and oxidative stress protection in membranes. *J. Mol. Biol.* 360, 934–945. doi: 10.1016/j.jmb.2006.05.056

Hughes, A. L. (1994). The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.* 256, 119–124. doi: 10.1098/rspb.1994.0058

Innan, H., and Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* 11, 97–108. doi: 10.1038/nrg2689

Jain, M., Ghanashyam, C., and Bhattacharjee, A. (2010). Comprehensive expression analysis suggests overlapping and specific roles of rice glutathione S-transferase genes during development and stress responses. *BMC Genomics* 11:73. doi: 10.1186/1471-2164-11-73

Jha, B., Sharma, A., and Mishra, A. (2011). Expression of SbGSTU (tau class glutathione S-transferase) gene isolated from *Salicornia* brachiata in tobacco for salt tolerance. *Mol. Biol. Rep.* 38, 4823–4832. doi: 10.1007/s11033-010-0625-x

Karavangeli, M., Labrou, N. E., Clonis, Y. D., and Tsaftaris, A. (2005). Development of transgenic tobacco plants overexpressing maize glutathione S-transferase I for chloroacetanilide herbicides phytoremediation. *Biomol. Eng.* 22, 121–128. doi: 10.1016/j.bioeng.2005.03.001

Kitamura, S., Shikazono, N., and Tanaka, A. (2004). TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in *Arabidopsis*. *Plant J.* 37, 104–114. doi: 10.1046/j.1365-313X.2003.01943.x

Koch, M. A., and Kiefer, M. (2005). Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species–*Capsella rubella*, *Arabidopsis lyrata* subsp. petraea, and A. thaliana. *Am. J. Bot.* 92, 761–767. doi: 10.3732/ajb.92.4.761

Kram, B. W., Xu, W. W., and Carter, C. J. (2009). Uncovering the *Arabidopsis thaliana* nectary transcriptome: investigation of differential gene expression in floral nectariferous tissues. *BMC Plant Biol.* 9:92. doi: 10.1186/1471-2229-9-92

Kwon, S. Y., Choi, S. M., Ahn, Y. O., Lee, H. S., Lee, H. B., Park, Y. M., et al. (2003). Enhanced stress-tolerance of transgenic tobacco plants expressing a human dehydroascorbate reductase gene. *J. Plant Physiol.* 160, 347–353. doi: 10.1078/0176-1617-00926

Lan, T., Wang, X. R., and Zeng, Q. Y. (2013). Structural and functional evolution of positively selected sites in pine glutathione S-transferase enzyme family. *J. Biol. Chem.* 288, 24441–24451. doi: 10.1074/jbc.M113.456863

Lan, T., Yang, Z. L., Yang, X., Liu, Y. J., Wang, X. R., and Zeng, Q. Y. (2009). Extensive functional diversification of the Populus glutathione S-transferase supergene family. *Plant Cell* 21, 3749–3766. doi: 10.1105/tpc.109.070219

Lieberherr, D., Wagner, U., Dubuis, P. H., Metraux, J. P., and Mauch, F. (2003). The rapid induction of glutathione S-transferases AtGSTF2 and AtGSTF6 by avirulent *Pseudomonas syringae* is the result of combined salicylic acid and ethylene signaling. *Plant Cell Physiol.* 44, 750–757. doi: 10.1093/pcp/pcg093

Liu, H. J., Tang, Z. X., Han, X. M., Yang, Z. L., Zhang, F. M., Yang, H. L., et al. (2015). Divergence in enzymatic activities in the soybean gst supergene family provides new insight into the evolutionary dynamics of whole-genome duplicates. *Mol. Biol. Evol.* 32, 2844–2859. doi: 10.1093/molbev/msv156

Liu, Y. J., Han, X. M., Ren, L. L., Yang, H. L., and Zeng, Q. Y. (2013). Functional divergence of the glutathione S-transferase supergene family in *Physcomitrella patens* reveals complex patterns of large gene family evolution in land plants. *Plant Physiol.* 161, 773–786. doi: 10.1104/pp.112.205815

Loyall, L., Uchida, K., Braun, S., Furuya, M., and Frohnmeyer, H. (2000). Glutathione and a UV light-induced glutathione S-transferase are involved in signaling to chalcone synthase in cell cultures. *Plant Cell* 12, 1939–1950. doi: 10.2307/3871204

Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., et al. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39, D225–D229. doi: 10.1093/nar/gkq1189

Marques, A. C., Vinckenbosch, N., Brawand, D., and Kaessmann, H. (2008). Functional diversification of duplicate genes through subcellular adaptation of encoded proteins. *Genome Biol.* 9, R54. doi: 10.1186/gb-2008-9-3-r54

Marrs, K. A. (1996). The functions and regulation of glutathione s-transferases in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 127–158. doi: 10.1146/annurev.arplant.47.1.127

Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W. (2003). Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* 15, 809–834. doi: 10.1105/tpc.009308

Moore, R. C., and Purugganan, M. D. (2005). The evolutionary dynamics of plant duplicate genes. *Curr. Opin. Plant Biol.* 8, 122–128. doi: 10.1016/j.pbi.2004.12.001

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

Nix, D. A., and Eisen, M. B. (2005). GATA: a graphic alignment tool for comparative sequence analysis. *BMC Bioinformatics* 6:9. doi: 10.1186/1471-2105-6-9

Oakley, A. J. (2005). Glutathione transferases: new functions. *Curr. Opin. Struct. Biol.* 15, 716–723. doi: 10.1016/j.sbi.2005.10.005

Ohno, S. (1970). *Evolution by gene duplication*. Heidelberg: Springer-Verlag.

Qian, W., and Zhang, J. (2009). Protein subcellular relocalization in the evolution of yeast singleton and duplicate genes. *Genome Biol. Evol.* 1, 198–204. doi: 10.1093/gbe/evp021

Ricci, G., Caccuri, A. M., Lo Bello, M., Pastore, A., Piemonte, F., and Federici, G. (1994). Colorimetric and fluorometric assays of glutathione transferase based on 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole. *Anal. Biochem.* 218, 463–465. doi: 10.1006/abio.1994.1209

Roxas, V. P., Smith, R. K. Jr., Allen, E. R., and Allen, R. D. (1997). Overexpression of glutathione S-transferase/glutathione peroxidase enhances the growth of transgenic tobacco seedlings during stress. *Nat. Biotechnol.* 15, 988–991. doi: 10.1038/nbt1097-988

Sappl, P. G., Carroll, A. J., Clifton, R., Lister, R., Whelan, J., Harvey Millar, A., et al. (2009). The *Arabidopsis* glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. *Plant J.* 58, 53–68. doi: 10.1111/j.1365-313X.2008.03761.x

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., et al. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nat. Genet.* 37, 501–506. doi: 10.1038/ng1543

Sharma, R., Sahoo, A., Devendran, R., and Jain, M. (2014). Over-expression of a rice tau class glutathione s-transferase gene improves tolerance to salinity and oxidative stresses in *Arabidopsis*. *PLoS ONE* 9:e92900. doi: 10.1371/journal.pone.0092900

Slotte, T., Hazzouri, K. M., Agren, J. A., Koenig, D., Maumus, F., Guo, Y. L., et al. (2013). The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nat. Genet.* 45, 831–835. doi: 10.1038/ng.2669

Soranzo, N., Sari Gorla, M., Mizzi, L., De Toma, G., and Frova, C. (2004). Organisation and structural evolution of the rice glutathione S-transferase gene family. *Mol. Genet. Genomics* 271, 511–521. doi: 10.1007/s00438-004-1006-8

Takahashi, Y., Hasezawa, S., Kusaba, M., and Nagata, T. (1995). Expression of the auxin-regulated parA gene in transgenic tobacco and nuclear localization of its gene products. *Planta* 196, 111–117. doi: 10.1007/BF00193224

Thom, R., Dixon, D. P., Edwards, R., Cole, D. J., and Lapthorn, A. J. (2001). The structure of a zeta class glutathione S-transferase from *Arabidopsis thaliana*: characterisation of a GST with novel active-site architecture and a putative role in tyrosine catabolism. *J. Mol. Biol.* 308, 949–962. doi: 10.1006/jmbi.2001.4638

Ushimaru, T., Nakagawa, T., Fujioka, Y., Daicho, K., Naito, M., Yamauchi, Y., et al. (2006). Transgenic *Arabidopsis* plants expressing the rice dehydroascorbate reductase gene are resistant to salt stress. *J. Plant Physiol.* 163, 1179–1184. doi: 10.1016/j.jplph.2005.10.002

Veitia, R. A., Bottani, S., and Birchler, J. A. (2008). Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects. *Trends Genet.* 24, 390–397. doi: 10.1016/j.tig.2008.05.005

Vickers, T. J., Wyllie, S., and Fairlamb, A. H. (2004). Leishmania major elongation factor 1B complex has trypanothione S-transferase and peroxidase activity. *J. Biol. Chem.* 279, 49003–49009. doi: 10.1074/jbc.M311039200

Walsh, B. (2003). Population-genetic models of the fates of duplicate genes. *Genetica* 118, 279–294. doi: 10.1023/A:1024194802441

Wang, X., Huang, Y., Lavrov, D. V., and Gu, X. (2009). Comparative study of human mitochondrial proteome reveals extensive protein subcellular relocalization after gene duplications. *BMC Evol. Biol.* 9:275. doi: 10.1186/1471-2148-9-275

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* 2:e718. doi: 10.1371/journal.pone.0000718

Xu, J., Tian, Y. S., Xing, X. J., Peng, R. H., Zhu, B., Gao, J. J., et al. (2015). Over-expression of AtGSTU19 provides tolerance to salt, drought and methyl viologen stresses in *Arabidopsis*. *Physiol. Plant.* 156, 164–175. doi: 10.1111/ppl.12347

Zhang, J. Z. (2003). Evolution by gene duplication: an update. *Trends Ecol. Evol.* 18, 292–298. doi: 10.1016/s0169-5347(03)00033-8

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 He, Guan, Chen, Gou, Liu, Zeng and Lan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 29 July 2016
doi: 10.3389/fpls.2016.01074

Nitrogen-Efficient and Nitrogen-Inefficient Indian Mustard Showed Differential Expression Pattern of Proteins in Response to Elevated CO_2 and Low Nitrogen

*Peerzada Y. Yousuf[1], Arshid H. Ganie[1], Ishrat Khan[1], Mohammad I. Qureshi[2], Mohamed M. Ibrahim[3,4], Maryam Sarwat[5], Muhammad Iqbal[1] and Altaf Ahmad[6]**

[1] Department of Botany, Faculty of Science, Jamia Hamdard, New Delhi, India, [2] Proteomics and Bioinformatics Laboratory, Department of Biotechnology, Faculty of Natural Sciences, Jamia Millia Islamia, New Delhi, India, [3] Department of Botany and Microbiology, Science College, King Saud University, Riyadh, Saudi Arabia, [4] Department of Botany and Microbiology, Faculty of Science, Alexandria University, Alexandria, Egypt, [5] Pharmaceutic Biotechnology, Amity Institute of Pharmacy, Amity University, Noida, India, [6] Proteomics and Nanobiotechnology Laboratory, Department of Botany, Faculty of Life Sciences, Aligarh Muslim University, Aligarh, India

OPEN ACCESS

Edited by:
Naser A. Anjum, University of Aveiro, Portugal

Reviewed by:
Jürgen Kreuzwieser, University of Freiburg, Germany
Ahmad Ali, University of Mumbai, India

***Correspondence:**
Altaf Ahmad
aahmad.bo@amu.ac.in

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *16 January 2016*
Accepted: *07 July 2016*
Published: *29 July 2016*

Citation:
Yousuf PY, Ganie AH, Khan I, Qureshi MI, Ibrahim MM, Sarwat M, Iqbal M and Ahmad A (2016) Nitrogen-Efficient and Nitrogen-Inefficient Indian Mustard Showed Differential Expression Pattern of Proteins in Response to Elevated CO_2 and Low Nitrogen. Front. Plant Sci. 7:1074. doi: 10.3389/fpls.2016.01074

Carbon (C) and nitrogen (N) are two essential elements that influence plant growth and development. The C and N metabolic pathways influence each other to affect gene expression, but little is known about which genes are regulated by interaction between C and N or the mechanisms by which the pathways interact. In the present investigation, proteome analysis of N-efficient and N-inefficient Indian mustard, grown under varied combinations of low-N, sufficient-N, ambient [CO_2], and elevated [CO_2] was carried out to identify proteins and the encoding genes of the interactions between C and N. Two-dimensional gel electrophoresis (2-DE) revealed 158 candidate protein spots. Among these, 72 spots were identified by matrix-assisted laser desorption ionization-time of flight/time of flight mass spectrometry (MALDI-TOF/TOF). The identified proteins are related to various molecular processes including photosynthesis, energy metabolism, protein synthesis, transport and degradation, signal transduction, nitrogen metabolism and defense to oxidative, water and heat stresses. Identification of proteins like PII-like protein, cyclophilin, elongation factor-TU, oxygen-evolving enhancer protein and rubisco activase offers a peculiar overview of changes elicited by elevated [CO_2], providing clues about how N-efficient cultivar of Indian mustard adapt to low N supply under elevated [CO_2] conditions. This study provides new insights and novel information for a better understanding of adaptive responses to elevated [CO_2] under N deficiency in Indian mustard.

Keywords: ***Brassica juncea*, proteomics, elevated CO_2, nitrogen efficiency, 2-DE**

Abbreviations: *B. juncea*, *Brassica juncea*; cv, cultivars; 2-DE, Two-dimensional gel electrophoresis; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; N, nitrogen; NE, nitrogen efficiency; NUE, nitrogen use efficiency.

INTRODUCTION

Carbon dioxide (CO_2), the main substrate for photosynthesis, plays a crucial role in growth, development, and productivity of plants. High CO_2 levels enhance the carboxylase activity and inhibit oxygenase activity of Rubisco, slowing down photorespiration. Increased CO_2 concentrations are expected to result in enhanced photosynthetic production of carbohydrates and other organic compounds. However, the photosynthetic efficiency relies not only on mere presence of CO_2 but also on assimilation, which is affected by the nitrogen (N) status of plant (Aranjueloa et al., 2014). At elevated [CO_2], low supply of N in the soil could limit the leaf area for intercepting light, restrain the capacity of plants to fix CO_2 photosynthetically, or lead to diminishing plant N availability over time through long-term soil-plant C and N dynamics (Reich and Hobbie, 2013). A large increase in biomass accumulation under elevated [CO_2], often observed in short-term experiments, may not be sustained over the long term in natural systems, given the N limiting conditions that predominate in both unmanaged and managed vegetations (Oren et al., 2001; Hungate et al., 2003; Luo et al., 2006; Reich et al., 2006). Given that limitations to productivity resulting from insufficient availability of N are widespread under elevated [CO_2], increased N supply is required for enhancing crop productivity in this condition. However, increased use of nitrogenous fertilizers is neither environmentally nor economically favorable. These fertilizers are too costly for poor farmers and their application also causes pollution of nitrate. Nitrogen-limiting conditions, therefore, reduce the responsiveness of plants to elevated [CO_2] and decrease the photosynthetic rate (Sanz-Saez et al., 2010). In order to sustain productivity under changing environmental factors, there is a need to look for plants that can utilize the positive effect of CO_2, even at low N.

Stress response genes can be figured out by expression profiling of plants, following the exposure to high levels of stress that can identify signaling components and their downstream effectors (Ahuja et al., 2010; Mitler et al., 2012; Hancock et al., 2014). Abiotic stress experiments impose a high stress level to identify processes and genes involved in plant survival under extreme conditions (Fowler and Thomashow, 2002; Umezawa et al., 2006; Jamil et al., 2011). Despite these triumphs, examples of such translational research to crop species are few (De Block et al., 2005; Li et al., 2008). Proteomics serves as the finest tool for investigating environmental pressures, genetic manipulation, stress-adaptive responses, and genotypic variability (Yousuf et al., 2015). The proteomic approach based on two-dimensional electrophoresis coupled with mass spectrometry provides an indispensable means to assess qualitative and quantitative changes of the proteome.

Indian mustard [*Brassica juncea* (L.) Czern. Coss.] is an important agricultural crop, grown primarily for oil production. After oil extraction, seed residue is used as animal feed. High rates of N fertilizer are usually applied to this crop in order to obtain the maximum seed yield because of its low harvest index (Schjoerring et al., 1995). Several studies are available on the response of this plant to elevated concentrations of CO_2 at physiological and biochemical levels (Frick et al., 1994; Uprety and Mahalaxmi, 2000; Uprety et al., 2001; Qaderi and Reid, 2005; Qaderi et al., 2006; Ruhil et al., 2015), but no effort has been made for identification of proteins in Indian mustard during the response to elevated [CO_2] accompanied with low N, to figure out the regulatory network during this phase. The present investigation was, therefore, undertaken for proteome analysis of N-efficient and N-inefficient Indian mustard grown under varied combinations of N-deficiency, N-sufficiency, ambient [CO_2], and elevated [CO_2], in order to identify genes and processes regulated by interactions between C and N metabolisms.

MATERIALS AND METHODS

Plant Culture and Treatments

In an earlier study, we reported Pusa Bold and Pusa Jai Kisan cultivars of Indian mustard (*B. juncea* L. Czern. Coss.) as N-efficient and N-inefficient, respectively (Ahmad et al., 2008). The seeds of these cultivars were washed thoroughly with distilled water and then surface sterilized with freshly prepared 0.01% mercuric chloride. These were rinsed with sterile distilled water (2–3 times) before sown in pots containing mixture of sand and vermiculite (1:1). After 3 days of germination, seedlings of similar size were placed in the half strength Hoagland's solution (pH 5.8) containing (mM): 1.0 KH_2PO_4, 3.0 KNO_3, 1.0 $MgSO_4$, and 0.5 NaCl and (μM) 23.1 H_3BO_3, 4.6 $MnCl_2$, 0.38 $ZnSO_4$, 0.16 $CuSO_4$, 0.052 H_2MoO_4, and 44.8 $FeSO_4$ (as ferric sodium-EDTA complex) on perforated polystyrene floats. The experiment was conducted in a completely randomized design with three replications. The nutrient solution was bubbled with sterile air to provide sufficient O_2 and changed on alternate days. The plants were grown in glasshouse under controlled temperature (27°C), light (16-h photoperiods) and humidity (60%) for 35 days in four sets of different treatment conditions. The glasshouse was divided in two chambers fitted with carbon dioxide gas cylinders along with the control system to maintain different CO_2 levels. In one set, plants were gown under sufficient-nitrogen (10 mM N) and ambient CO_2 levels (T0, Control). Second set of plants were grown at low-nitrogen (1 mM N) and ambient CO_2 levels (T1). In the third set, sufficient-nitrogen (10 mM) and elevated CO_2 (500 ppm) levels were given to plants (T2). Low nitrogen (1 mM) and elevated CO_2 (500 ppm) levels were maintained for the fourth set of plants (T3). Nitrogen was supplied in the form of nitrate (KNO_3) in all the sets. Leaves of 35-day-old plants were sampled, immediately dipped in liquid nitrogen, and stored at −80°C till the proteomic analysis was carried out.

Protein Extraction

Proteins were extracted from leaf samples using the phenol method of Isaacson et al. (2006), wherein 2 g of leaf material was ground to fine powder in liquid nitrogen and suspended in 10 ml of extraction buffer containing 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), β-mercaptoethanol, sucrose, and phenylmethanesulfonylfluoride (PMSF). Fifteen milliliters of phenol was added to this solution, mixed in a cold room rocker for 30 min and subjected to centrifugation at 5000 rpm for 10 min at 4°C. The top phenolic phase was carefully

recovered in a separate tube and incubated at −20°C overnight for precipitation after adding 15 ml of ice-cold 0.1 M ammonium acetate solution. The proteins were pelleted by centrifuging at 10,000 rpm for 15 min at 4°C. The pellet was washed once with methanol and then twice with chilled acetone. The resulting pellet was centrifuged at 5000 rpm after each washing and then dried and solubilised in the buffer containing 2 M thiourea, 7 M urea, 4% CHAPS, 50 mM DTT. The protein was quantified using the Bradford reagent (Bio-Rad, USA).

Two-Dimensional Gel Electrophoresis

Two-dimensional electrophoresis of proteins was performed to resolve the leaf proteome. In the first dimensional run, IPG strips (24 cm, pH 3–10, NL; Bio-Rad, USA) were used. From each treatment 500 μg protein (in 400 μl rehydration buffer) was loaded through passive rehydration at 20°C for 14 h. Isoelectric focussing was carried out in a PROTEAN IEF apparatus (Bio-Rad, USA). The voltages applied were 250 V for 1 h, 500 V for 1 h, 1000 V for 2 h, 2000 V for 2 h, linear increase of 8000 V and running till achieving 80,000 Vh, followed by a slow ramping of 500 V for 1 h. After the completion of IEF, the strips were exposed to reduction buffer for 15 min and then to alkylation buffer for 15 min. The SDS-PAGE was carried out in a Dodeca cell (Bio-Rad PROTEAN plus, USA) for separation of focussed proteins, using 12% SDS at a constant voltage of 250 V. The gels were stained with colloidal Coommassie brilliant blue dye and then destained with ultrapure water.

Gel Analysis

The resolved gels were scanned by densitometer (GS-800 Calibrated Densitometer, Bio-Rad, USA) and then analyzed with the help of PD Quest software (Advanced version 8.0 Bio-Rad, Hercules, CA, USA) for spot detection, background subtraction, and intensity quantification. 2D maps from all treatments were compared for spot number and retrieving relative volume of each and every spot against the control gel. The normalization of each spot value was done in terms of percentage of the total volume of all gel spots for rectification of unevenness due to quantitative disparity in spot intensities. The spots with two-fold change in their volumes during the treatment or a significant variation between the control and other treatments as per the results of paired Student's *t*-test ($p \leq 0.05$) were spotted as the treatment-responsive proteins. Such proteins were selected and further analyzed for their identification through mass spectrometry.

In-Gel Digestion and Protein Identification

The protein spots were excised from gels, washed and dehydrated with acetonitrile and ammonium bicarbonate and then reduced with 15 mM DTT at 60°C for 1 h. The gel slices were alkylated by 100 mM isoamyl alcohol in the dark for 15 min, rehydrated with ammonium bicarbonate and then dried up in a speed vac for 15 min. The dried gel slices were subjected to rehydration with 15 μl of working trypsin (Sequencing grade, Promega, USA) at 37°C overnight. The supernatant was taken and 20% acetonitrile and 1% formic acid were added to the remaining gel slice for further extraction. The final supernatant was dried in speed vac until the volume was reduced to 25–50 μl. This final volume was analyzed with AB Sciex MALDI-TOF MS, as mentioned in Bagheri et al. (2015). Peptide tolerance of 150 ppm, fragment mass tolerance of ±0.4 Da, and peptide charge of 1+ were selected. Classical protein database searches were performed on a local Mascot (Matrix Science, London, UK) server. Only significant hits, as defined by the MASCOT probability analysis ($p < 0.05$), were accepted. Peptides were searched with the following parameters: NCBInr database, taxonomy of green plants, trypsin of the digestion enzyme, one missed cleavage site, partial modification of cysteine carboamidomethylated, and methionine oxidized. In addition, the searches were performed without constraining protein Mr and pI.

Statistical Analyses

Three biological replicates for each of the treatments and control were used for statistical analyses. A two-tailed Students *t*-test with the significance of 95% was performed on the normalized value of protein spots with the help of SPSS software.

RESULTS

Response of *Brassica juncea* Proteomes to Elevated CO_2 and Low-N Conditions

Comparative proteomics of leaves from both cultivars of *B. juncea*, following exposure to elevated [CO_2] and low nitrogen helped in unraveling interesting proteins. Representative gels from the control and treatment plants are shown in **Figure 1**. In total, around 452 protein spots were visualized in each gel falling between the pH range of 3–10. Detailed information about the proteome changes was obtained by scanning the gels that were digitized using PD Quest™ (Advanced version 8.0 Bio-Rad, Hercules, CA, USA). 2D maps of both the mustard cultivars grown under control and treatment conditions were compared. Of the spots visualized, 72 showed more than two-fold change in abundance by the treatments. However, the pattern of differential expression of these protein spots varied in both the cultivars under treatments of N and CO_2 (**Table 1**). Under the ambient CO_2 level, the number of differentially expressed proteins was 21 in Pusa Jai Kisan and 11 in Pusa Bold when comparison was made between low-N and sufficient-N conditions**.** Under elevated [CO_2], the differentially expressed protein spots were 27 and 13 in Pusa Jai Kisan and Pusa Bold, respectively, when comparison was made between low-N and sufficient-N conditions. Interestingly, differentially expressed protein spots under the condition of low-N treatment were 10 and 21 in Pusa Jai Kisan and Pusa Bold, respectively, when a comparison was made between the treatments of ambient and elevated CO_2 levels. Under the same CO_2 treatments, the number of differentially expressed protein spots was similar (19) in both the cultivars with the supply of sufficient-N.

Analysis of Differentially Expressed Protein Spots

MALDI-TOF MS enabled the identification of 72 differentially expressed protein spots (**Table 2**, Supplementary Table). Out of these differentially expressed proteins, 48 (67%) were upregulated and 24 (33%) downregulated. These were localized to different positions on 2-DE plots (**Figure 2**). Of the 72 identified proteins, 67 exhibited homology with proteins of known function. Based

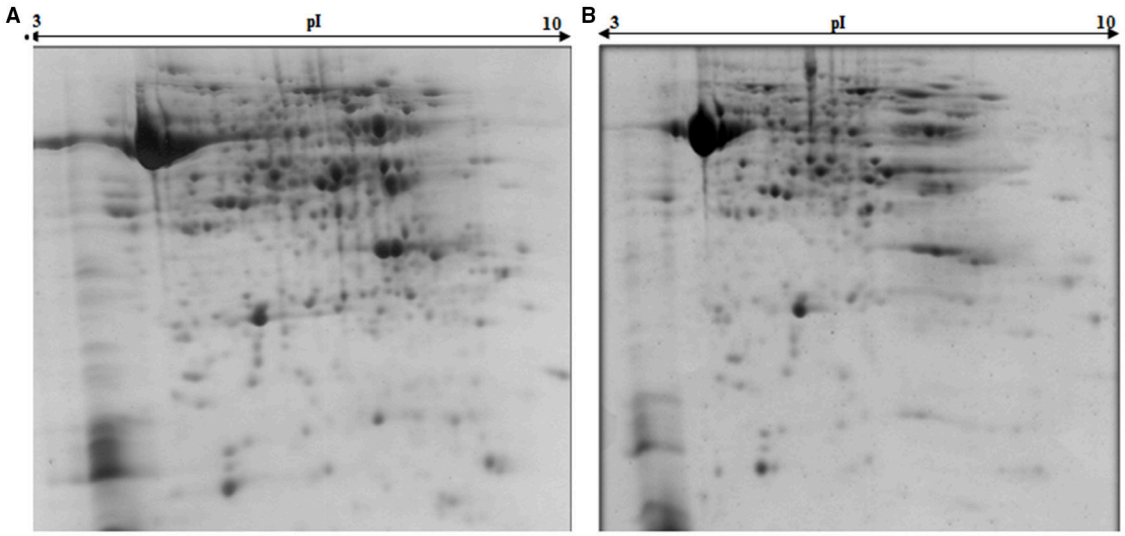


FIGURE 1 | 2DE maps of leaf samples representing Indian mustard cultivars, (A) cv. Pusa Bold and (B) Pusa Jai Kisan.

TABLE 1 | Differential expression of protein spots in Pusa Jai Kisan (N-inefficient) and Pusa Bold (N-efficient) cultivars of Indian mustard under the treatments of low-N, sufficient-N, ambient CO_2, and elevated CO_2 levels.

Treatments comparison	Cultivars of Indian mustard	
	Pusa Jai Kisan	Pusa Bold
T1/T0	21	11
T2/T3	27	13
T1/T3	10	21
T0/T2	19	19

T1/T0, ambient CO_2 level and low-N vs. ambient CO_2 level and sufficient-N; T2/T3, elevated CO_2 level and sufficient-N vs. elevated CO_2 level and low-N; T1/T3, ambient CO_2 level and low-N vs. elevated CO_2 level and low-N; T0/T2, ambient CO_2 level and sufficient-N vs. elevated CO_2 level and sufficient-N.

on the function and physiological processes involved in, these were categorized into 13 major groups, showing their association with biosynthesis (10%), carbohydrate metabolism (9%), energy metabolism (3%), photosynthesis (12%), protein folding (6%), transcription and signaling (6%), oxidative stress (10%), water stress (5%), lipid metabolism (6%), heat tolerance (3%), nitrogen metabolism (5%) and protein synthesis (10%), the rest of them (15%) were unclassified proteins. These proteins showed differential relative spot intensities not only between mustard cultivars but also under different treatments (**Figure 3**).

Spatial Categorization of Differentially Expressed Proteins

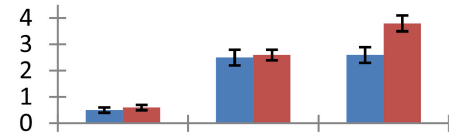
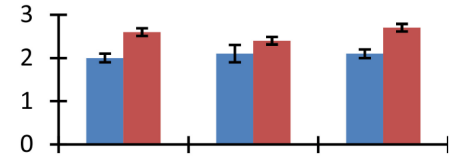
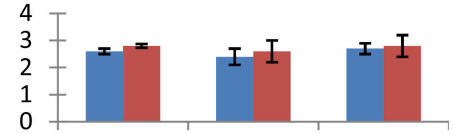
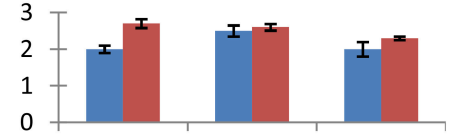
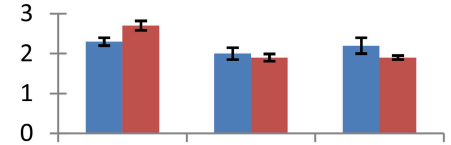
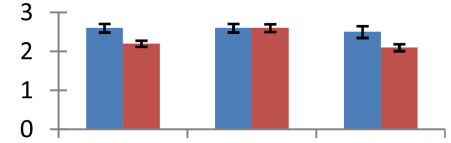
The differentially expressed proteins belonged to different sub-cellular sites (**Figure 4**). Proteins from almost all the cellular sites showed changes in abundance, indicating the effect of combinatorial action of elevated [CO_2] and low-N on the organelle functioning. The maximum number of proteins belonged to chloroplasts (32%), followed by cytosol (21%), nucleus (21%), and mitochondria (9%) while a relatively low number of them belonged to ribosomes (2%), Golgi bodies (2%), vacuoles (2%), endoplasmic reticulum (3%), and plasma membrane (6%).

DISCUSSION

Influence on Biosynthetic Pathways

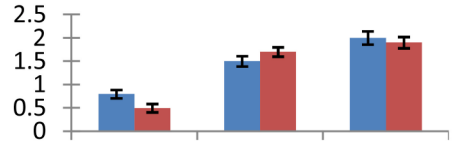
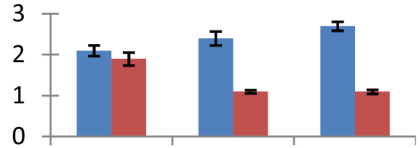
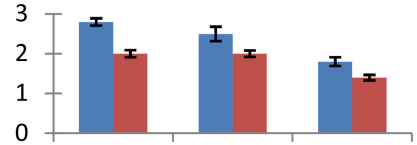
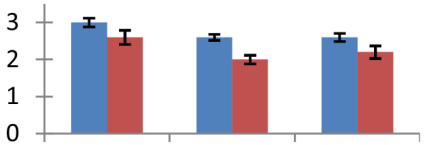
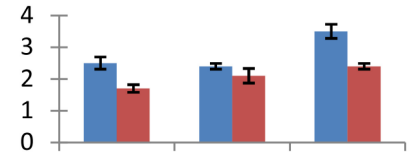
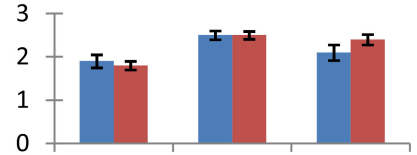
Combinatorial impact of elevated [CO_2] and low N altered the abundance of many proteins that are involved in synthesis of different biological components including amino acids, hormones, and various cellular structural components. Cuticle plays a vital role in retaining integrity of plant in changing environment. Ketoacyl-CoA synthase, an important fatty-acyl elongase, helps in elongation of acyl chains during synthesis of cuticular wax. This enzyme (spot 34) showed a significant increase in expression, indicating the need for sealing of plant surfaces during the treatment conditions. The maximum upregulation of protein occurred in elevated [CO_2] and low-N, showing the necessity of wax formation under these environmental conditions. Cultivar Pusa Bold showed a higher increase in expression levels than cv. Pusa Jai Kisan. Increase in thickness of cuticle wax has been observed earlier in plants grown during different abiotic stresses including high temperature, water deficit and high irradiation (Shepherd and Griffiths, 2006). Isopropylmalate dehydratase (spot 11), an enzyme with a vital role in leucine synthesis, showed a sizeable reduction in expression in treated samples with respect to control; the maximum reduction was observed in cv. Pusa Jai Kisan. The treatments reduced the expression of proteins involved in the synthesis of lignin (spot 57), phytochrome chromophore (spot 71), glycan (spot 66), ethylene (spot 47) in both the cultivars. NAD-dependent epimerase hydratase (spot 65), an enzyme with

TABLE 2 | Identification, sub-cellular localization, and quantitative analysis of differentially expressed proteins.

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
1 (U)	gi|111608879	α-1,4-glucan phosphorylase	8852	5.17	134	Chloroplast			
2 (U)	gi|226533870	cp31BHv	18,293	4.85	86	Nucleus, chloroplast			
3 (U)	gi|145327751	GLYOXYLASE I 7	15,564	5.22	74	Peroxisome			
4 (U)	gi|88909669	L-ascorbate peroxidase 5	33,852	5.84	94	Chloroplast			
5 (D)	gi|1174745	Triosephosphate isomerase, chloroplastic	31,955	6.10	148	Chloroplast, cytosol			
6 (U)	gi|259490070	Lipid binding protein precursor	12,904	6.69	59	Membrane			

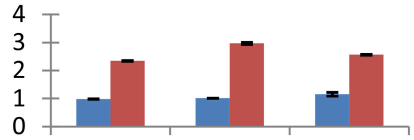
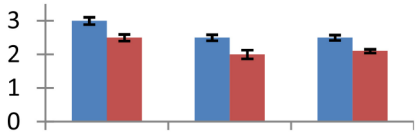
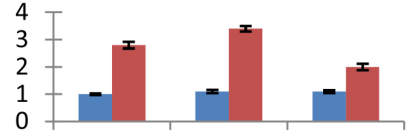
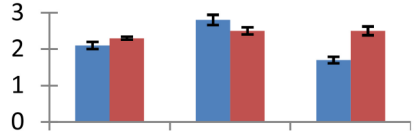
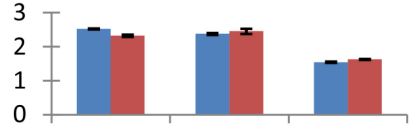
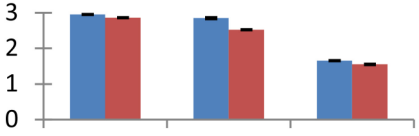
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities
							Low N effect / eCO_2 effect / eCO_2 + Low N effect
7 (U)	gi|6523032	Flavonol synthase-like protein	36,401	5.33	94	Cytoplasm, nucleus	
8 (U)	gi|15229664	MYB77; DNA binding/transcription factor	34,318	6.07	53	Nucleus	
9 (D)	gi|521301355	Ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit	53,901	6.16	351	Chloroplast	
10 (D)	gi|15220615	Chlorophyll a-b binding protein 1	24,836	5.11	112	Chloroplast	
11 (D)	gi|226493349	3-isopropylmalate dehydratase large subunit 2	53,901	7.05	147	Chloroplast	
12 (U)	gi|162135976	Sterolesin-B	28,873	5.63	55		

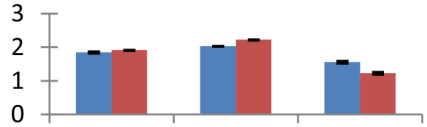
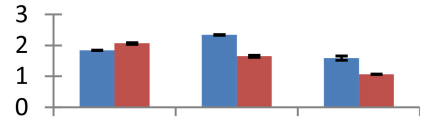
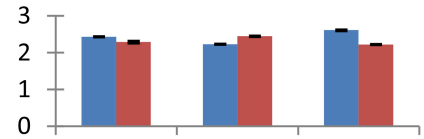
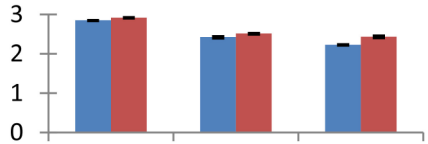
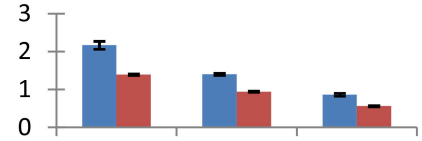
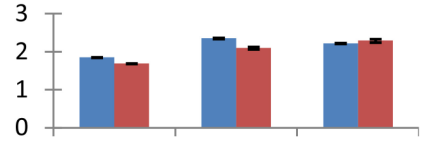
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
13 (U)	gi|37653227	PII-like protein	25,729	9.84	160	Chloroplast			
14 (U)	gi|21322655	ADP glucose pyrophosphatase	21,172	5.62	83	Chloroplast			
15 (U)	gi|334183671	Glutamine synthetase	46,852	5.96	120	Chloroplast/mitochondria			
16 (U)	gi|255073229	Glutathione transferase	25,098	6.35	76	Cytosol			
17 (U)	gi|297382831	Beta-amylase	24,276	5.59	64	Chloroplast, vacuole, cytosol			
18 (D)	gi|27735224	Ribulose bisphosphate carboxylase small chain 2B	20,801	6.29	232	Chloroplast			

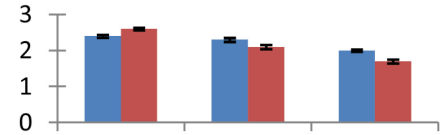
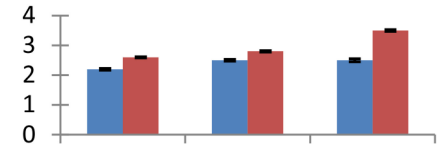
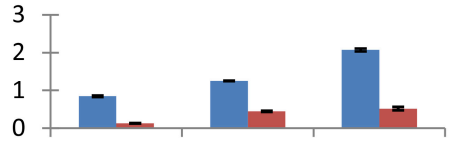
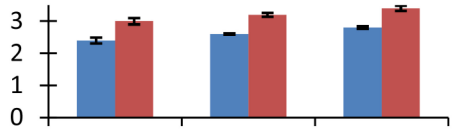
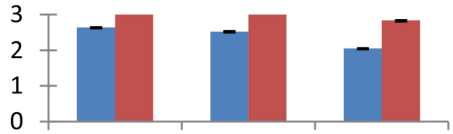
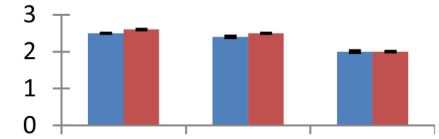
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities
							Low N effect / eCO_2 effect / eCO_2 + Low N effect
19 (U)	gi|224015956	RADIALIS	9381	8.29	32	Nucleus	
20 (U)	gi|31415978	Putative TNP2-like transposase	116,347	7.91	60	Nucleus	
21 (D)	gi|108706066	Cytoplasmic aconitate hydratase	99,314	5.73	151	Cytosol	
22 (U)	gi|159157513	Putative chaperonin 60 beta precursor	64,721	5.68	158	Mitochondria	
23 (U)	gi|2244775	Salt-inducible protein homolog	65,905	5.96	51	Nucleus	
24 (U)	gi|15231865	AGD6; ARF GTPase activator/DNA binding/zinc ion binding	49,222	6.57	99	Golgi apparatus	

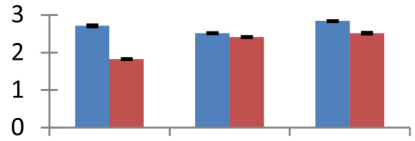
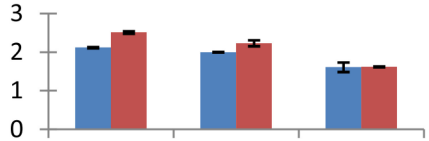
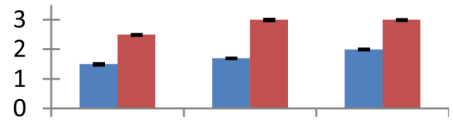
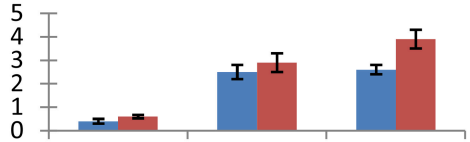
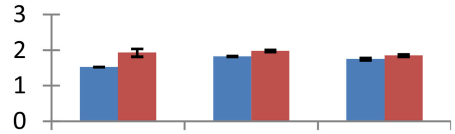
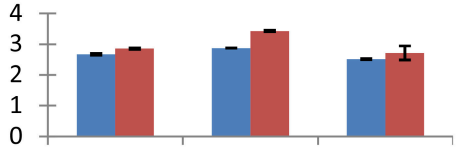
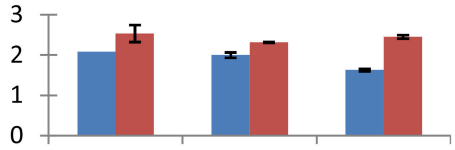
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
25 (D)	gi|49574618	Ribosomal protein L14	13,216	9.71	54	Ribosome			
26 (D)	gi|114329695	Ribosomal protein S19	10,689	10.90	50	Ribosome, chloroplast, mitochondria			
27 (U)	gi|325947633	Salt overly sensitive 2	20,331	4.94	33	Plasma membrane			
28 (U)	gi|100831	Photosystem II oxygen-evolving complex protein 1	33,596	8.13	160	Chloroplast			
29 (U)	gi|1263097	MYB-related protein	33,799	6.07	133	Nucleus			
30 (D)	gi|2546952	Translation elongation factor-TU	27,375	5.09	99	Plastid, mitochondria, cytosol			

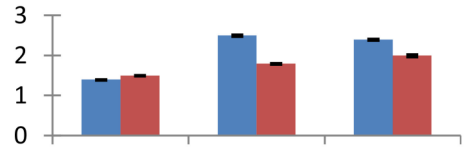
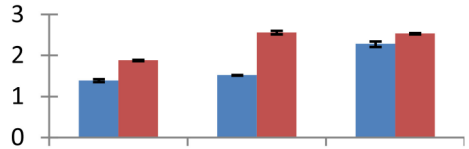
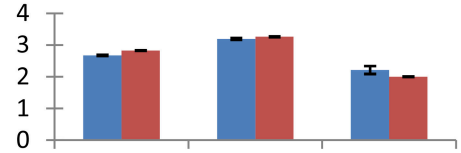
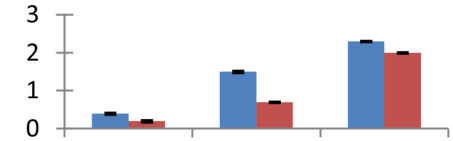
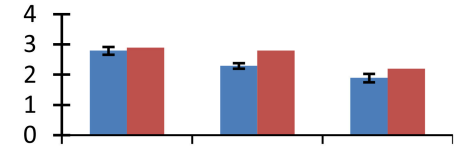
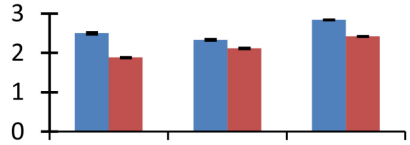
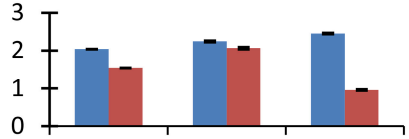
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
31 (U)	gi|4090259	Ubiquitin-conjugating enzyme E2	18,891	7.01	75	Nucleus, cytosol			
32 (U)	gi|145666464	Protein disulfide isomerase	56,921	5.01	149	ER			
33 (U)	gi|12643259	Rubisco activase chloroplast precursor	51,886	5.43	138	Chloroplast			
34 (U)	gi|330255651	3-Ketoacyl-CoA synthase	46,810	9.02	65	Plasma membrane			
35 (U)	gi|297807473	GDSL-motif lipase/hydrolase family protein	43,124	8.76	34	chloroplast stroma			
36 (U)	gi|1572627	Cu/Zn superoxide dismutase	20,312	5.35	130	Cytosol			
37 (U)	gi|15231902	Phosphatidylinositol-4-phosphate 5-kinase family protein	82,129	8.63	133	Plasma membrane			

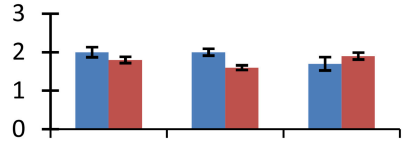
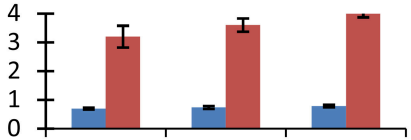
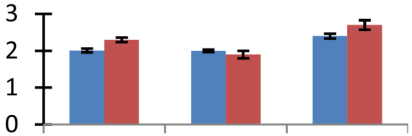
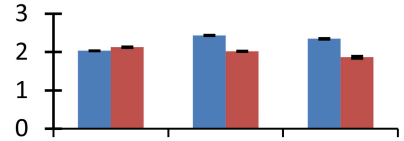
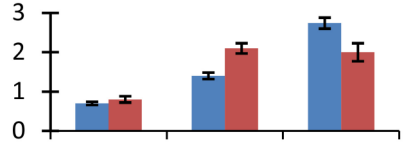
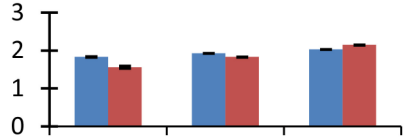
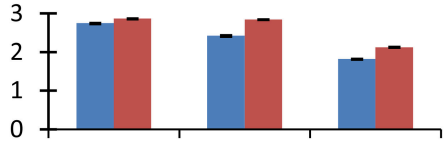
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities
							Low N effect / eCO_2 effect / eCO_2 + Low N effect
38 (U)	gi|30686609	PUB23 (PLANT U-BOX 23); ubiquitin-protein ligase	46,648	8.28	92	Cytosol	
39 (D)	gi| 15229231	Glyceraldehyde 3-phosphate dehydrogenase	37,366	6.62	141	Plastid, mitochondria	
40 (U)	gi|729477	RecName: Ferredoxin-NADP reductase	41,622	8.54	139	Chloroplast stroma, thylakoid membrane	
41 (D)	gi|296932709	Granule-bound starch synthase I	13,696	9.76	51	Chloroplast	
42 (U)	gi|50252278	Flavin-containing monooxygenase	42,795	8.75	65	ER	
43 (D)	gi|186886309	ATPase F1 alpha subunit	38,233	8.27	134	Mitochondria	
44 (D)	gi|260780700	Acetyl-CoA carboxylase beta subunit	40,007	9.02	53	Cytoplasm, plastid	

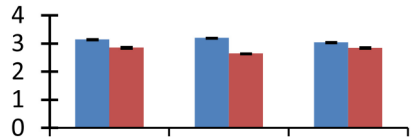
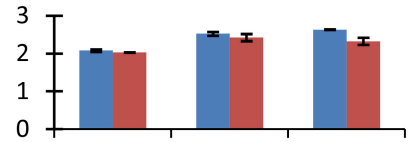
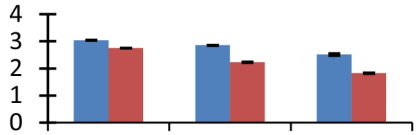
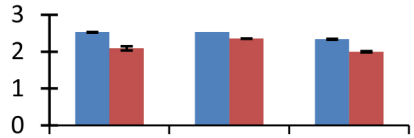
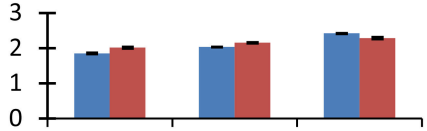
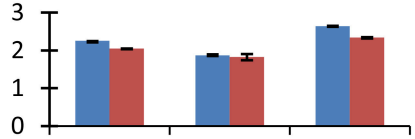
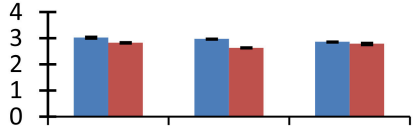
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
45 (U)	gi|2570499	23 kDa polypeptide of photosystem II	27,772	9.06	94	Chloroplast			
46 (U)	gi|83700338	Chloroplast-localized cyclophilin	16,127	8.48	146	Chloroplast			
47 (D)	gi|52353463	ACC oxidase	36,852	4.97	134	Cytosol			
48 (D)	gi|217075182	Unknown	19,372	8.81	39				
49 (U)	gi|145329965	TTN7 (TITAN7); ATP binding/protein binding	138,856	6.05	144	Nucleus			
50 (D)	gi|56785387	Hypothetical protein [*Oryza sativa* Japonica Group]	8515	11.31	40				
51 (U)	gi|21537409	Osmotin-like protein	27,148	6.34	135	Cell-wall			

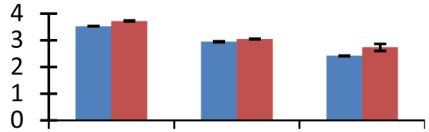
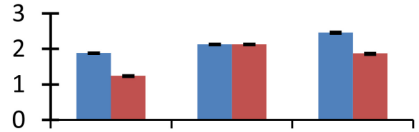
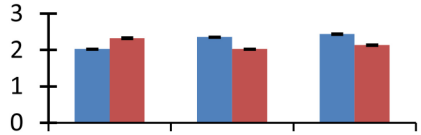
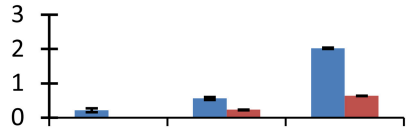
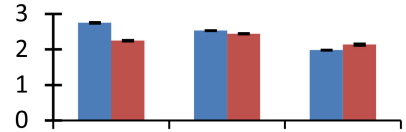
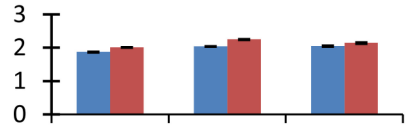
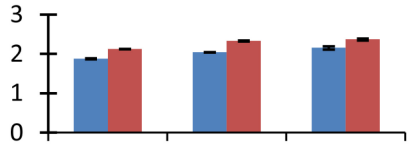
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
52 (D)	gi|22296371	Putative 29 kDa ribonucleoprotein A, chloroplast precursor	28,126	4.75	165	Nucleus			
53 (D)	gi|41584495	SMC3	139,879	6.07	54	Cytoplasm, nucleus			
54 (D)	gi|79595267	ATPase	209,868	6.50	169	Plasma membrane, tonoplast			
55 (U)	gi|255575181	Carnitine racemase	27,262	7.64	32	Nucleus, peroxisome			
56 (U)	gi|297807151	Hypothetical protein ARALYDRAFT_909075	27,949	5.57	45				
57 (D)	gi|260619528	LIM1	21,920	9.07	124	Nucleus, cytosol			
58 (D)	gi|69216892	RNA polymerase beta chain	79,308	9.10	217	Nucleus			

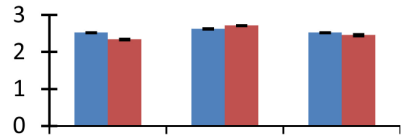
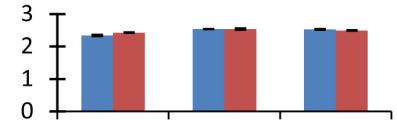
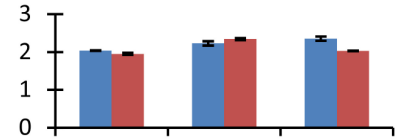
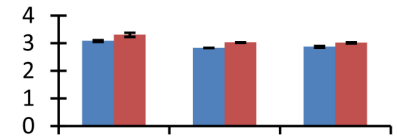
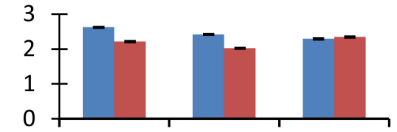
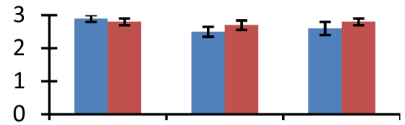
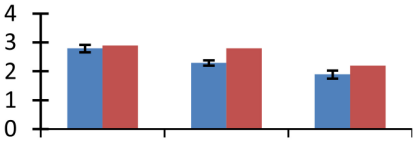
(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities		
							Low N effect	eCO_2 effect	eCO_2 + Low N effect
59 (U)	gi|21397266	Small heat shock protein	25,122	6.26	61	Nucleus			
60 (U)	gi|269969449	Iron deficiency-specific protein	374,514	5.71	78	Cytosol			
61 (U)	gi|38344419	OSJNBb0080H08.11	202,531	8.71	46				
62 (U)	gi|304571953	Anthranilic acid methyltransferase 3	43,177	6.29	142	Cytosol			
63 (D)	gi|4261643	Sedoheptulose-1,7-bisphosphatase	42,068	6.26	153	Chloroplast			
64 (U)	gi|18424254	Zinc finger (C3HC4-type RING finger) family protein	44,981	6.36	88	Nucleus			
65 (U)	gi|159487407	NAD-dependent epimerase/dehydratase	36,518	7.13	135	Cytosol			

(Continued)

TABLE 2 | Continued

S.N. (U/D)	Accession no.	Name of protein	Mr (da)	pI	M. Score	Location	Relative spot intensities
							Low N effect / eCO_2 effect / eCO_2 + Low N effect
66 (D)	gi|297806809	Galactosyl transferase GMA12/MNN10 family protein	52,488	7.52	95	Golgi apparatus	
67 (U)	gi|113205313	"Chromo" domain containing protein	93,199	9.21	143	Thylakoid membrane	
68 (U)	gi|15213029	Maturase K	14,517	9.86	55	Mitochondria, chloroplast, nucleus	
69 (U)	gi|308804756	Stress-induced protein sti1-like protein (ISS)	69,580	6.03	72	Cytosol, chloroplast	
70 (U)	gi|238005590	Unknown	9609	8.88	52		
71 (D)	gi|75180109	Heme oxygenase 4	32,115	6.89	62	Chloroplast	
72 (U)	gi|300669644	Ferredoxin-dependent glutamate synthase 2	27,542	6.24	102	Chloroplast	

S.N. (U/D), Spot Number; U, upregulated; D, downregulated; Mr, Molecular Weight; pI, Isoelectric Point; M.Sco, MASCOT Score; eCO_2, elevated CO_2 level. Y-axis denoted relative spot intensities; X-axis denotes treatments. Red bars indicate Pusa Bold cultivar, and blue bars indicate Pusa Jai Kisan cultivar. Spot volumes were analyzed by PD Quest software. The fold-change of up-regulated protein spot volumes was calculated by treatment (T1-T3)/control (T0), whereas the change fold of down-regulated protein spot volumes was calculated by control (T0)/treatment (T1-T3). T0, ambient CO_2 level and sufficient-N; T1, ambient CO_2 level and low-N; T2, elevated CO_2 level and sufficient-N; T3, low-N and elevated CO_2 level. Values are presented as mean of three replicates ± SE.

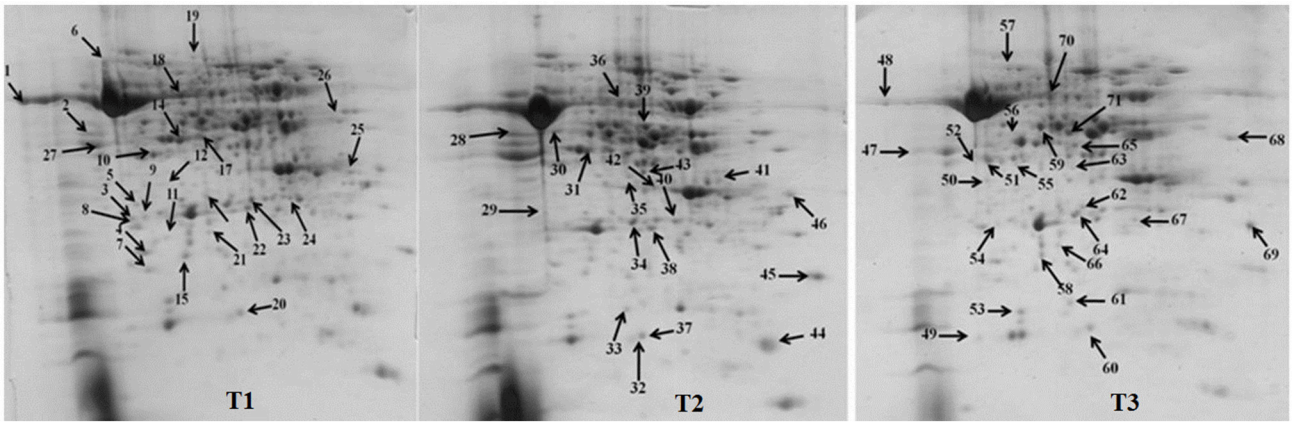


FIGURE 2 | Position of differentially expressed protein spots on 2DE map of Indian mustard. The differentially expressed proteins illustrated on three different gel images (T1, T2, T3 treatments) were those showing two-fold changes in their relative abundance and hence selected for mass spectrometric analysis for identification.

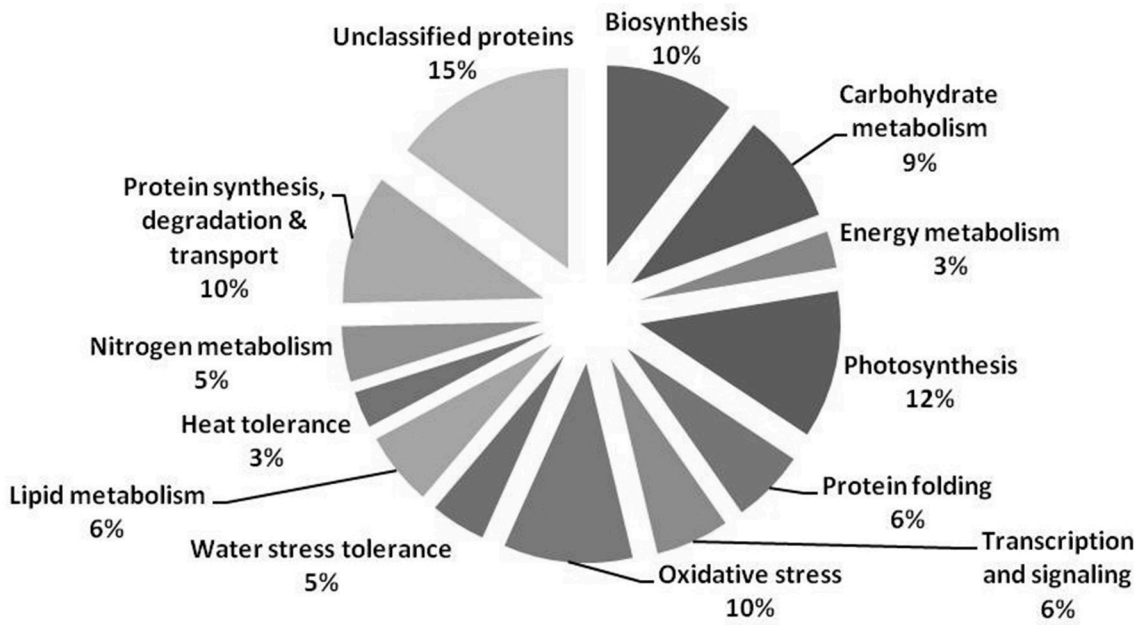


FIGURE 3 | Pie diagram showing the functional cataloguing of differentially expressed proteins.

a crucial role in biosynthesis of auxin, was upregulated in both the cultivars. Pusa Bold exhibited a higher level of this protein than Pusa Jai Kisan.

Photosynthesis

Combined stress of elevated [CO_2] and low nitrogen altered expression of proteins regulating different aspects of photosynthesis including photoxidation of water, electron transport, and carbon fixation. Chlorophyll a/b binding protein 1 forms an integral part of light harvesting complex, which captures light and delivers excitation energy to the photosystem. Treatment conditions resulted in a sizeable down-regulation of this protein (spot 10) signifying the effect of these conditions on light captivity and photoinhibition. Oxygen-evolving enhancer protein-1 is essential for the normal functioning of PSII and plays a critical role in the stabilization of Mn cluster *in vivo* (Yi et al., 2005), besides regulating the turnover of D1 protein of PSII reaction center (Lundin et al., 2007). A considerable increase in the expression of this protein (spot 28) was monitored in treated samples. Cultivar Pusa Bold was more responsive than cv. Pusa Jai Kisan, exhibiting its competence to withstand the negative effects of treatment conditions on the PS II

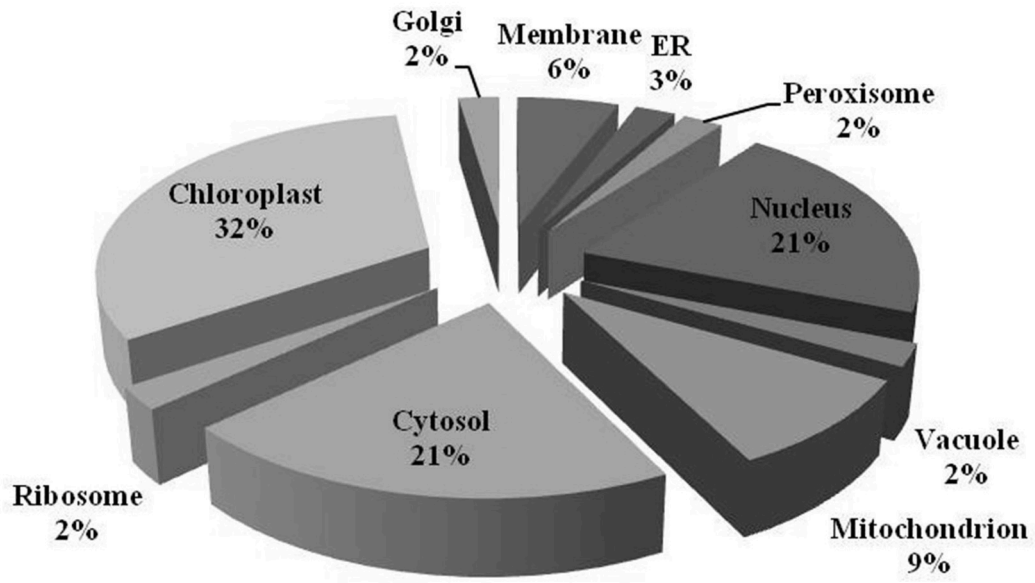


FIGURE 4 | Pie diagram depicting the spatial distribution of differentially expressed proteins at subcellular level.

functioning. The overexpression of oxygen-evolving complex provides tolerance to plants under different stresses including the osmotic, salinity, and heavy metals stress (Gururania et al., 2013).

The expression of 23 KDa polypeptide of PS II, a protein having an indispensable role in the restoration of oxygen evolution activity by generating a high-affinity binding site for Ca^{2+} on the oxidizing side of photosystem II, was upregulated (spot 45) in both the mustard cultivars under the treatments of elevated [CO_2] and low-N. Rubisco, the primary CO_2-fixing enzyme in plants, was downregulated during treatments with respect to the control (spot 9 and 18). The degree of reduction was more intense under low N conditions than other two treatments, signifying the effect of N deficiency on C fixation. The degradation and decrease in rubisco levels offers an excellent opportunity in plants to liberate amino acids that are in turn reutilized to regulate N level (Feller et al., 2008) under N-deficient conditions. Rubisco activase, an important molecular chaperone that acts as rubisco conformational switch, activating the enzyme from inactive state (Spreitzer and Salvucci, 2002), showed a considerable increase (spot 33) during the treatment in both the mustard cultivars, compared to their respective controls. Cultivar Pusa Bold exhibited higher level of expression of this protein than cv. Pusa Jai Kisan, confirming its higher efficiency to maintain optimal photosynthtic rate under the given combined stress. Although photosynthetic rate increases during the individual effect of high [CO_2] level (Taub, 2010), the observed down-regulation of enzymes associated with photosynthetic efficiency indicated the masking of this effect under combined stress.

Carbohydrate Metabolism

The treatments induced an adverse effect on the carbohydrate metabolic pathway as indicated by the downregulation of enzymes involved in glycolysis (Triose isomerase, spot 5, and glyceraldehyde 3-phosphate dehydrogenase, spot 39) and C3 cycle including cytoplasmic aconitate hydratase (spot 21) and sedoheptulose-1,7-bisphosphatase (spot 63), and starch biosynthesis (granule-bound starch synthase I, spot 41). The reduced abundance of enzymes involved in glycolysis and Calvin cycle may be attributed to the decreased rate of carbon fixation induced by low nitrogen. Beta-amylase, which catalyzes degradation of starch, glycogen, and related polysaccharides to produce beta-maltose, was upregulated (spot 17). ADP-glucose pyrophosphorylase (spot 14), an enzyme regulating the inhibition of starch and APP-glucose syhthesis was upregulated during the treatment, causing adverse effects on carbohydrate synthesis.

Protein Folding

Environmental stress changes functional conformation and stabilization of proteins. To overcome this problem, plants possess specific proteins that facilitate folding and shield other proteins from aggregation and misfolding. Four such proteins, cp31BHv (spot 2), chaperonin 60 beta precursor (spot 22), PDI (spot 32), and cyclophilin (spot 46) were identified that showed varied levels of expression in both of the cultivars, Pusa Bold having the relatively higher ones. Of the four proteins, cyclophilins, encompassing molecular chaperones with scaffolding, foldase and chaperoning properties, are of great impotance as they regulate a number of metabolic pathways and

perform diverse functions in plants (Kumari et al., 2013). High expression levels of cyclophilin gene family have been observed in plants under various environmental stresses including salinity, drought, cold, and heat (Trivedi et al., 2013).

Heat Shock Tolerance

Elevated [CO_2] causes increase in temperature due to enhanced greenhouse effect, which leads to heat stress, inducing denaturation of proteins. The heat shock proteins (HSPs) maintain stability of proteins during heat stress, besides regulating various cellular processes, including those associated with tolerance to multiple environmental stresses (Song et al., 2014). The small HSP (spot 59) exhibited higher expression in treated materials, compared to control, with a higher intensity in cv. Pusa Bold than cv. Pusa Jai Kisan. Zinc finger (C3HC4-type RING finger) family protein (spot 64), another heat-inducible protein that helps in combating drought stress, also showed higher levels of expression under conditions of low-N and elevated [CO_2] in cv. Pusa Bold than in cv. Pusa Jai Kisan, thus showing the heat tolerance capacity of the former cultivar.

Oxidative Stress

Almost all environmental stresses cause overproduction of reactive oxygen species (ROS) that lead to oxidative stress in tissues. In order to scavenge the toxic ROS, plants possess sophisticated antioxidant defense system. Nine proteins that are part of antioxidant system showed differential expression. Among these proteins, APX (spot 4), glutathione transferase (spot 16), Cu/Zn SOD (spot 36), and Fd-NADP reductase (spot 40) are directly involved in glutathione-ascorbate pathway. Flavonol-synthase like protein (spot 7), which catalyses production of flavonoids, is involved in oxidative defense, besides other processes including auxin transport regulation, protection during UV stress and cell signaling (Harborne and Williams, 2000). Sterolesin B (spot 12), glyoxylase (spot 3), and flavin-containing monooxygenase (spot 42) are also known to perform important defense functions during the oxidative stress (Nicolas et al., 2007). All these oxidative defense proteins were upregulated under treatment conditions and the expression was higher in cv. Pusa Bold than cv. Pusa Jai Kisan, signifying its greater efficiency to scavenge the toxic ROS.

Nitrogen Metabolism

Nitrogen metabolism of field crops carries greatest significance with reference to their nutritional status (Cui et al., 2009; Yue et al., 2012). Three proteins with critical roles in N metabolism showed diverged expression pattern. PII-like protein (spot 13), involved in nitrogen sensing in *Arabidopsis* (Hsieh et al., 1998) and rice (Sugiyama et al., 2004), showed upsurge in expression. However, cv. Pusa Bold experienced a two to three-fold increase in expression of this protein. We have recorded a higher efficiency of N-efficient cultivar (cv. Pusa Bold) in sensing and uptake of nitrogen under N-limited conditions than the nitrogen-inefficient cultivar (cv. Pusa Jai Kisan). Glutamine synthetase (GS), an important enzyme of N assimilation, catalyzes production of ammonia generated from different processes including nitrate and ammonia metabolism, nitrogen fixation, photorespiration and catabolism of proteins and compounds meant for nitrogen transport (Miflin and Habash, 2002). This ammonia assimilatory protein (spot 15) accumulated in both cultivars but spot intensity was higher in cv. Pusa Bold than cv. Pusa Jai Kisan. The GS gene overexpresses in response to different abiotic stresses (Cai et al., 2009). Fd-dependent glutamate synthase is an important enzyme in plants that helps in ammonium assimilation through GS/GOGAT pathway (Reitzer, 2003). This enzyme couples with GS to catalyze incorporation of ammonia in 2-oxoglutarate. Overexpression of glutamate synthase (spot 72) was observed in both the cultivars.

Lipid Metabolism

The level of expression of four proteins involved in lipid metabolism, carnitine racemase (spot 55), GDSL-motif lipase/hydrolase family protein (spot 35), acetyl-CoA carboxylase (spot 44), and lipid-binding protein precursor (spot 6) was altered. While the former two are involved in lipid catabolism, acetyl-CoA carboxylase is involved in the biosynthesis of fatty-acids and the last one in lipid transport (Hancock et al., 2014). GDSL-motif lipase/hydrolase family proteins are thought to play an imperative role in morphogenesis and plant development (Akoh et al., 2004). Carnitine racemase helps in β-oxidation of fats. The spot intensity of proteins involved in lipid catabolism was increased while that of proteins involved in fatty-acid synthesis was decreased in both cultivars under the experimental conditions, as compared with the control. The increased levels of carnitine racemase implies shortage of glucose and increased energy demand under the treatment conditions.

Protein Synthesis, Degradation, and Transport

Protein synthesis was inhibited under stress, as evidenced by down-regulation of many proteins associated with the protein synthesis machinery, such as protein L14 (spot 25), S19 (spot 26), 29 KDa ribonucleoprotein (spot 52), and translation elongation factor (spot 30). The decrease in expression of these proteins was steep in both cultivars. The low N availability may possibly be the reason for the reduced protein synthesis. Ubiquitination regulates degradation as well as localization of proteins besides other processes like transcriptional activation and protein-protein interactions (Xu et al., 2009). The two proteins associated with ubiquitination pathway, viz. PUB23 ubiquitin-protein ligase (spot 38) and ubiquitin-conjugating enzyme E2 (spot 31), were spotted to undergo up-regulation. Cultivar Pusa jai Kisan exhibited higher accumulation than Pusa Bold, indicating its sensitivity to protein degradation. ADP-ribosylation factor GTPase-activating protein AGD6 is an important enzyme, which facilitates protein trafficking to multiple organelles. Considerable upregulation of AGD6 (spot 24) was observed in mustard during stressful condition in contrast to the control.

Water Stress

α-1,4-glucan phosphorylase is an important enzyme that brings about phosphorolysis of terminal residues of α-1,4-linked glucan chains from non-reducing ends, generating glucan-1-phosphate as the product. The enzyme plays a vital role during transient

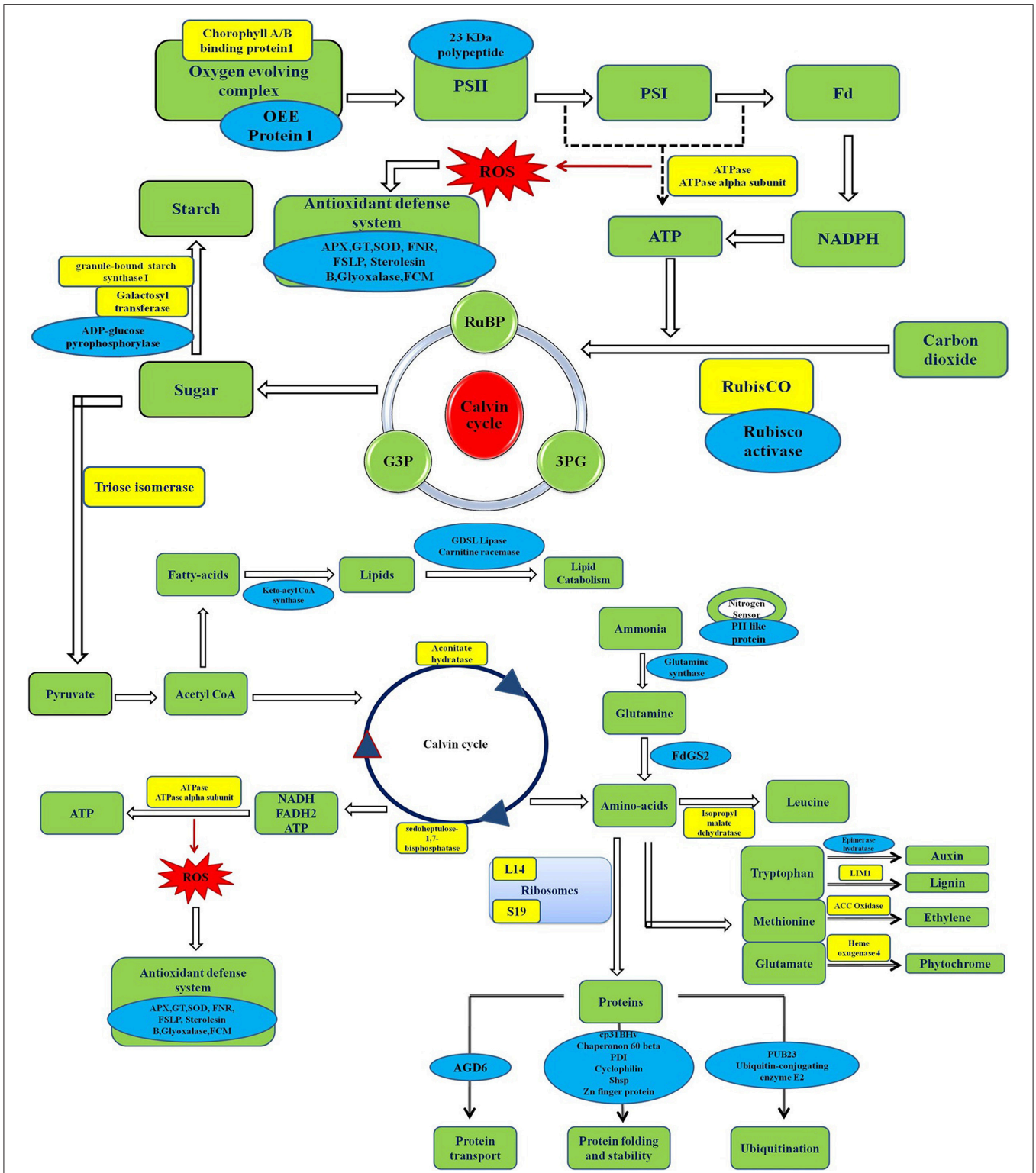


FIGURE 5 | Schematic representation of differentially-expressed proteins in Indian mustard grown under teratments of [CO_2] and N. Proteins in yellow boxes were downregulated whereas those in blues boxes were upregulated. OEE protein1, Oxygen evolving enhance protein1; RUBP, Ribulose-1,5-bisphosphate; 3PG, 3-phosphoglycerate; G3P, Glyceraldehyde 3-phosphate; FdGS2, Ferredoxin-dependent glutamate synthase 2; AGD6, AGD6 ARF GTPase activator/DNA binding/zinc ion binding; GT, Glutathione transferase; FNR, Fd-NADP reductase; FCM, Flavin-containing monooxygenase.

water stress by supplying substrates for the respiratory metabolic reactions in the chloroplast (Zeeman et al., 2004). α-1,4-glucan phosphorylase (spot 1) exhibited higher expression level in cv. Pusa Bold than in cv. Pusa Jai Kisan under elevated CO_2 and low-N treatment. Osmotin like protein, known to play a crucial role in osmotic adjustment of plant cells besides other functions like disease resistance, among others (Anzlovar and Dermastia, 2003), upregulated under treatment conditions (spot 51), as compared with the control, showing a greater intensity in cv. Pusa Bold than in cv. Pusa Jai Kisan. The increased expression of the proteins regulating the synthesis of osmoprotectants may possibly be a means to overcome the negative effects of water stress induced by the elevated [CO_2].

ATP Synthesis

The treatment conditions reduced ATP synthesis as evidenced by down-regulation of two protein spots (43 and 54) related to the machinery responsible for energy production. The decrease in intensity of these proteins was, more intense in cv. Pusa Jai Kisan than cv. Pusa Bold. The reduced rate of ATP synthesis might be due to decline in the photsynthetic rate.

Transcription and Signaling

Treatment conditions caused a significant decrease in RNA polymerase beta chain (spot 58) which affected the transcription rate. The degree of down-regulation was more in cv. Pusa Jai Kisan than in cv. Pusa Bold. MYB77 (spot 8) and MYB-related protein (29), which act as transcription factors modulating auxin (Shin et al., 2007) abscisic acid (Abe et al., 1997) signal transduction, was upregulated under treatment conditions, with a high expression rate of MYB77 in cv. Pusa Jai Kisan and of MYB-related protein in cv. Pusa Bold. Phosphatidylinositol-4-phosphate 5-kinase family protein phosphorylates phosphatidylinositol-4-phosphate, produces phosphatidylinositol-4,5-bisphosphate, which acts as a precursor of two secondary messengers, namely phosphatidylinositol-4,5-bisphosphate and inositol-1,4,5-triphosphate. The treatments induced significant upregulation in this protein (spot 37) and the level of expression was higher in cv. Pusa Bold than in cv. Pusa Jai Kisan.

Unclassified Proteins

Ten proteins with diverse functions expressed differentially. Of these, SOS2 (spot 27) and salt-inducible protein homolog (spot 23) were associated with salt stress. These proteins are overexpressed in plants during exposure to salt. Iron deficiency has been detected in various food crops under the effect of high CO_2 concentrations (Myers et al., 2014). Iron deficiency-specific protein (spot 60) got upregulated under treatment conditions.

CONCLUSION

In response to elevated [CO_2] and low N treatments, many proteins involved in nitrate and C assimilation pathways were expressed differentially (**Figure 5**). Proteins that are involved in photosynthesis reactivation and in maintenance of chloroplast functionality exhibited change in expression pattern under elevated [CO_2] and low N conditions. Proteins associated with defense mechanism against heat, water, low nitrogen, and oxidative stresses upregulated and the degree of expression was higher in cv. Pusa Bold than in cv. Pusa Jai Kisan. Majority of the proteins related to biosynthesis of cellular components, photosynthesis, carbohydrate anabolism, and ATP synthesis were down-regulated. Major changes in protein expression pattern were observed in N-efficient (Pusa Bold) cultivar of mustard, showing its ability to grow well under elevated [CO_2] and low N. These results underline the strict relationship between N and C metabolisms. Five proteins, namely cyclophilin, elongation factor-TU, PII-like protein, oxygen-evolving complex I, and rubisco activase, may be considered by plant breeders and biotechnologists as suitable candidates for developing cultivars suitable to grow in conditions of elevated [CO_2] and low N availability without any penalty on productivity.

AUTHOR CONTRIBUTIONS

PY, IK, AA, MQ conceived and designed the experiments. PY and IK performed the experiments. PY, MS, AG, AA analyzed the data. PY, AA, MQ, MI, MMI wrote and revised the paper.

ACKNOWLEDGMENTS

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for support through research group no. RGP-297.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.01074

REFERENCES

Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D., and Shinozaki, K. (1997). Role of *Arabidopsis* MYC *and* MYB homologs in drought and abscisic acid-regulated gene expression. *Plant Cell* 9, 1859–1868.

Ahmad, A., Khan, I., Abrol, Y. P., and Iqbal, M. (2008). Genotypes variation in nitrogen use efficiency of mustard. *Environ. Pollut.* 154, 462–466. doi: 10.1016/j.envpol.2007.10.007

Ahuja, I., de Vos, R. C. H., Bones, A. M., and Hall, R. D. (2010). Plant molecular stress responses face climate change. *Trends Plant Sci.* 15, 664–674. doi: 10.1016/j.tplants.2010.08.002

Akoh, C. C., Lee, G., Liaw, Y., Huang, T., and Shaw, J. (2004). GDSL family of serine esterases/lipases. *Prog. Lipid Res.* 43, 534–552. doi: 10.1016/j.plipres.2004.09.002

Anzlovar, S., and Dermastia, M. (2003). The comparative analysis of osmotins and osmotin-like PR-5 proteins. *Plant Biol.* 5, 116–124. doi: 10.1055/s-2003-40723

Aranjueloa, I., Cabrerizoa, P. M., Aparicio-Tejoa, P. M., and Arrese-Igor, C. (2014). Unravelling the mechanisms that improve photosynthetic performance of N_2-fixing pea plants exposed to elevated [CO_2]. *Environ. Exp. Bot.* 99, 167–174. doi: 10.1016/j.envexpbot.2013.10.020

Bagheri, R., Bashir, H., Ahmad, J., Iqbal, M. and Qureshi, M. I. (2015). Spinach (*Spinacia olerecea* L.) modulates its proteome differentially in response to salinity, cadmium and their combination stress. *Plant Physiol. Biochem.* 97, 235–245. doi: 10.1016/j.plaphy.2015.10.012

Cai, H., Zhou, Y., Xiao, J., Li, X., Zhang, Q., and Lian, X. (2009). Overexpressed glutamine synthetase gene modifies nitrogen metabolism and abiotic stress responses in rice. *Plant Cell Rep.* 28, 527–537. doi: 10.1007/s00299-008-0665-z

Cui, Z., Zhang, F., Dou, Z., Yuxin, M., Sun, Q., Chen, X., et al. (2009). Regional evaluation of critical nitrogen concentrations in winter wheat production of the North China Plain. *Agron. J.* 101, 159–166. doi: 10.2134/agronj2008.0102

De Block, M., Verduyn, C., De Brouwer, D., and Cornelissen, M. (2005). Poly (ADPribose) polymerase in plants affects energy homeostasis, cell death and stress tolerance. *Plant J.* 41, 95–106. doi: 10.1111/j.1365-313X.2004.02277.x

Feller, U., Anders, I., and Demirevska, K. (2008). Degradation of rubisco and other chloroplast proteins under abiotic stress. *Gen. Appl. Plant Physiol.* 34, 5–18.

Fowler, S., and Thomashow, M. (2002). Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *Plant Cell* 14, 1675–1690. doi: 10.1105/tpc.003483

Frick, J., Nielsen, S., and Mitchell, C. A. (1994). Yield and seed oil content response of dwarf, rapid-cycling Brassica to nitrogen treatments, planting density, and carbon dioxide enrichment. *J. Am. Soc. Hortic. Sci.* 119, 1137–1143.

Gururania, M. A., Upadhyayab, C. P., Strasserc, R. J., Yua, J. W., and Parka, S. W. (2013). Evaluation of abiotic stress tolerance in transgenic potato plants with reduced expression of PSII manganese stabilizing protein. *Plant Sci.* 19, 7–16. doi: 10.1016/j.plantsci.2012.09.014

Hancock, R. D., Morris, W. L., Ducreux, L. J., Morris, J. A., Usman, M., Verrall, S. R., et al. (2014). Physiological, biochemical and molecular responses of the potato (*Solanum tuberosum* L.) plant to moderately elevated temperature. *Plant Cell Environ.* 37, 439–450. doi: 10.1111/pce.12168

Harborne, J. B., and Williams, C. A. (2000). Advances in flavonoid research since 1992. *Phytochemistry* 55, 481–504. doi: 10.1016/S0031-9422(00)00235-1

Hsieh, M. H., Lam, H. M., Loo, F. J. V. D., and Coruzzi, G. (1998). A PII-like protein in *Arabidopsis*: putative role in nitrogen sensing. *Proc. Nat. Acad. Sci. U.S.A.* 95, 13965–13970. doi: 10.1073/pnas.95.23.13965

Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y. Q., and Field, C. B. (2003). Atmospheric science. Nitrogen and climate change. *Science* 302, 1512–1513. doi: 10.1126/science.1091390

Isaacson, T., Damasceno, C. M, Saravanan, R. S., He, Y., Catala, C., Saladie, M., et al. (2006). Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nat. Protocols* 1, 769–774. doi: 10.1038/nprot.2006.102

Jamil, A., Riaz, S., Ashraf, M., and Foolad, M. R. (2011). Gene expression profiling of plants under salt stress. *Cri. Rev. Plant Sci.* 30, 435–458. doi: 10.1080/07352689.2011.605739

Kumari, S., Roy, S., Singh, P., Pareek, S. L., and Pareek, A. (2013). Cyclophilins proteins in search of function. *Plant Signal. Behav.* 8, 25–32. doi: 10.4161/psb.22734

Li, H., Payne, W. A., Michels, G. J., and Charles, M. R. (2008). Reducing plant abiotic and biotic stress: drought and attacks of greenbugs, corn leaf aphids and virus disease in dryland sorghum. *Environ. Exp. Bot.* 63, 305–316. doi: 10.1016/j.envexpbot.2007.11.014

Lundin, B., Hansson, M., Schoefs, B., Vener, A. V., and Spetea, C. (2007). The *Arabidopsis* $PsbO_2$ protein regulates dephosphorylation and turnover of the photosystem II reaction centre D1 protein. *Plant J.* 49, 528–539. doi: 10.1111/j.1365-313X.2006.02976.x

Luo, J., Sun, S., Jia, L., Chen, W., and Shen, Q. (2006). The mechanism of nitrate accumulation in pakchoi [*Brassica campestris* L. ssp. *Chinensis (L.)]. Plant Soil* 282, 291–300. doi: 10.1007/s11104-005-6094-7

Miflin, B. J., and Habash, D. Z. (2002). The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. *J. Exp. Bot.* 53, 979–987. doi: 10.1093/jexbot/53.370.979

Mitler, R., Finka, A., and Goloubinoff, P. (2012). How do plants feel the heat? *Trends Biochem. Sci.* 37, 118–125. doi: 10.1016/j.tibs.2011.11.007

Myers, S. S., Zanobetti, A., Kloog, I., Huybers, P., Leakey, A. D., Bloom, A. J., et al. (2014). Increasing CO_2 threatens human nutrition. *Lett. Nat.* 510, 139–143. doi: 10.1038/nature13179

Nicolas, S. D., LeMignon, G., Eber, F., Coriton, O., Monod, H., and Clouet, V. (2007). Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids. *Genetics* 175, 487–503. doi: 10.1534/genetics.106.062968

Oren, R., Ellsworth, D. S., Johnsen, K. H., Phillips, N., Ewers, B. E., Maier, C., et al. (2001). Soil fertility limits carbon sequestration by forest ecosystems in a CO_2-enriched atmosphere. *Nature* 411, 469–472. doi: 10.1038/35078064

Qaderi, M. M., Kurepin, L. V., and Reid, D. M. (2006). Growth and physiological responses of canola (*Brassica napus*) to three components of global climate change: temperature, carbon dioxide and drought. *Physiol. Plant.* 128, 710–721. doi: 10.1111/j.1399-3054.2006.00804.x

Qaderi, M. M., and Reid, D. M. (2005). Growth and physiological responses of canola (*Brassica napus*) to UV-B and CO_2 under controlled environment conditions. *Physiol. Plant.* 125, 247–259. doi: 10.1111/j.1399-3054.2005.00566.x

Reich, P. B., Hobbie, S. E., Lee, T., Ellsworth, D. S., West, J. B., Tilman, D., et al. (2006). Nitrogen limitation constrains sustainability of ecosystem response to CO2. *Nature* 440, 922–925. doi: 10.1038/nature04486

Reich, P. B., and Hobbie, S. E. (2013). Decade-long soil nitrogen constraint on the CO_2fertilization of plant biomass. *Nat. Clim. Change* 3, 278–282. doi: 10.1038/nclimate1694

Reitzer, L. (2003). Nitrogen assimilation and global regulation in *Escherichia coli*. *Annu. Rev. Microbiol.* 57, 155–176. doi: 10.1146/annurev.micro.57.030502.090820

Ruhil, K., Sheeba, Ahmad, A., Iqbal, M., and Tripathy, B. C. (2015). Photosynthesis and growth responses of mustard (*Brassica juncea* L. Cv. Pusa Bold) plants to free air carbon dioxide enrichment (FACE). *Protoplasma* 252, 935–946. doi: 10.1007/s00709-014-0723-z

Sanz-Saez, A., Erice, G., Aranjuelo, I., Nogues, S., Irigoyen, J. J., and Sanchez-Diaz, M. (2010). Photosynthetic down-regulation under elevated CO_2 exposure can be prevented by nitrogen supply in nodulated alfalfa. *J. Plant Physiol.* 167, 1558–1565. doi: 10.1016/j.jplph.2010.06.015

Schjoerring, J. K., Bock, J. G. H., Gammelvind, L., Jensen, C. R., and Mogensen, V. O. (1995). Nitrogen incorporation and remobilization in different shoot components of field-grown winter oilseed rape (*Brassica napus* L.) as affected by rate of nitrogen application and irrigation. *Plant Soil* 177, 255–264.

Shepherd, T., and Griffiths, D. W. (2006). The effects of stress on plant cuticular waxes. *New Phytol.* 171, 469–499. doi: 10.1111/j.1469-8137.2006.01826.x

Shin, R., Burch, A. Y., Huppert, K. A., Tiwari, S. B., Murphy, A. S., Guilfoyle, T. J., et al. (2007). The *Arabidopsis* transcription factor MYB77 modulates auxin signal transduction. *Plant Cell* 19, 2440–2453. doi: 10.1105/tpc.107.050963

Song, A., Zhu, X., Chen, F., Gao, H., Jiang, J., and Chen, S. (2014). A *Chrysanthemum* heat shock protein confers tolerance to abiotic stress. *Int. J. Mol. Sci.* 15, 5063–5078. doi: 10.3390/ijms15035063

Spreitzer, R. J., and Salvucci, M. E. (2002). Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.* 53, 449–475. doi: 10.1146/annurev.arplant.53.100301.135233

Sugiyama, K., Hayakawa, T., Kudo, T., Ito, T., and Yamaya, T. (2004). Interaction of N-acetylglutamate kinase with a PII-Like protein in rice. *Plant Cell Physiol.* 45, 1768–1778. doi: 10.1093/pcp/pch199

Taub, D. R. (2010). Effects of rising atmospheric concentrations of carbon dioxide on plants. *Nat. Educ. Knowl.* 1, 21.

Trivedi, D. K., Ansari, M. W., and Tuteja, N. (2013). Multiple abiotic stress responsive rice Cyclophilin (OsCYP-25) mediates a wide range of cellular responses. *Commun. Integ. Biol.* 6, 1–8. doi: 10.4161/cib.25260

Umezawa, T., Fujita, M., Fujita, Y., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2006). Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. *Curr. Opin. Chem. Biol.* 17, 113–122. doi: 10.1016/j.copbio.2006.02.002

Uprety, D. C., Dwivedi, N., Mohan, R., and Paswan, G. (2001). Effect of elevated CO_2 concentration on leaf structure of *Brassica juncea* under water stress. *Biol. Plant.* 44, 149–152. doi: 10.1023/A:1017959429783

Uprety, D. C., and Mahalaxmi, V. (2000). Effect of elevated CO_2 and nitrogen nutrition on photosynthesis, growth and carbon-nitrogen balance in *Brassica juncea*. *J. Agron. Crop Sci.* 184, 271–276. doi: 10.1046/j.1439-037x.2000.00392.x

Xu, L., Menard, R., Berr, A., Fuchs, J., Cognat, V., Meyer, D., et al. (2009). The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of FLC expression and repression of flowering in *Arabidopsis thaliana*. *Plant J.* 57, 279–288. doi: 10.1111/j.1365-313X.2008.03684.x

Yi, X., McChargue, M., Laborde, S., Frankel, L. K., and Bricker, T. M. (2005). The manganese-stabilizing protein is required for photosystem II assembly/stability and photoautotrophy in higher plants. *J. Biol. Chem.* 280, 16170–16174. doi: 10.1074/jbc.M501550200

Yousuf, P. Y., Ahmad, A., Ganie, A. H., and Iqbal, M. (2015). Salt stress-induced modulations in the shoot proteome of *Brassica juncea* genotypes. *Environ. Sci. Pollut. Res.* 23, 2391–2401. doi: 10.1007/s11356-015-5441-3

Yue, S., Meng, Q., Zhao, R., Li, F., Chen, X., Zhang, F., et al. (2012). Critical nitrogen dilution curve for optimizing nitrogen management of winter wheat production in the North China Plain. *Agron. J.* 104, 523–529. doi: 10.2134/agronj2011.0258

Zeeman, C. S., Thorneycroft, D., Schupp, N., Chapple, A., Weck, M., Dunstan, H., et al. (2004). Plastidial a-glucan phosphorylase is not required for starch degradation in *Arabidopsis* leaves but has a role in the tolerance of abiotic Stress. *Plant Physiol.* 135, 849–858. doi: 10.1104/pp.103.032631

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Yousuf, Ganie, Khan, Qureshi, Ibrahim, Sarwat, Iqbal and Ahmad. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 April 2016
doi: 10.3389/fpls.2016.00541

Interactions of Sulfate with Other Nutrients As Revealed by H_2S Fumigation of Chinese Cabbage

*Martin Reich[1], Muhammad Shahbaz[2], Dharmendra H. Prajapati[1], Saroj Parmar[3], Malcolm J. Hawkesford[3] and Luit J. De Kok[1]**

[1] Laboratory of Plant Physiology, Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, Netherlands, [2] Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester, MA, USA, [3] Plant Biology and Crop Science Department, Rothamsted Research, Harpenden, UK

Sulfur deficiency in plants has severe impacts on both growth and nutrient composition. Fumigation with sub-lethal concentrations of H_2S facilitates the supply of reduced sulfur via the leaves while sulfate is depleted from the roots. This restores growth while sulfate levels in the plant tissue remain low. In the present study this system was used to reveal interactions of sulfur with other nutrients in the plant and to ascertain whether these changes are due to the absence or presence of sulfate or rather to changes in growth and organic sulfur. There was a complex reaction of the mineral composition to sulfur deficiency, however, the changes in content of many nutrients were prevented by H_2S fumigation. Under sulfur deficiency these nutrients accumulated on a fresh weight basis but were diluted on a dry weight basis, presumably due to a higher dry matter content. The pattern differed, however, between leaves and roots which led to changes in shoot to root partitioning. Only the potassium, molybdenum and zinc contents were strongly linked to the sulfate supply. Potassium was the only nutrient amongst those measured which showed a positive correlation with sulfur content in shoots, highlighting a role as a counter cation for sulfate during xylem loading and vacuolar storage in leaves. This was supported by an accumulation of potassium in roots of the sulfur-deprived plants. Molybdenum and zinc increased substantially under sulfur deficiency, which was only partly prevented by H_2S fumigation. While the causes of increased molybdenum under sulfur deficiency have been previously studied, the relation between sulfate and zinc uptake needs further clarification.

Keywords: Brassica, hydrogen sulfide, sulfur deficiency, yield quality, mineral composition

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Holger Hesse,
Freien Universität Berlin, Germany
Agata Gadaleta,
University of Bari, Italy

****Correspondence:***
Luit J. De Kok
l.j.de.kok@rug.nl

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *14 January 2016*
Accepted: *05 April 2016*
Published: *27 April 2016*

Citation:
Reich M, Shahbaz M, Prajapati DH, Parmar S, Hawkesford MJ and De Kok LJ (2016) Interactions of Sulfate with Other Nutrients As Revealed by H_2S Fumigation of Chinese Cabbage. Front. Plant Sci. 7:541. doi: 10.3389/fpls.2016.00541

INTRODUCTION

Understanding interactions between plant nutrients is essential for optimizing fertilization strategies and improving nutrient use efficiency in crops (Baxter, 2009; Maathuis, 2009; Reich et al., 2014a). Sulfur was recognized as an essential nutrient for crops more than a century ago (Bogdanov, 1899; Hart and Peterson, 1911) and since then it had shown to be involved in many vital processes in plants (Thompson, 1967; Hell, 1997; De Kok et al., 2005; Hawkesford and De Kok, 2006; Takahashi et al., 2011). In contrast to the intensive study of the uptake and assimilation

of sulfur (reviewed e.g., by Leustek et al., 2000; Kopriva, 2006; Hawkesford, 2012; Honsel et al., 2012) and their interconnection with carbon and nitrogen metabolism (Kopriva et al., 2002) the interaction with other nutrients is less well-known.

As sulfur deficiency is becoming a constraint to yield in many cropping systems throughout the world (Zhao et al., 1999) it is momentous to uncover the effects on other elements, which determine the nutritional quality of crops. Numerous members of the Brassicaceae family are used as food and oil crops worldwide and Chinese cabbage has an increasing importance in many developing countries (Kawashima and Soares, 2003; Rakow, 2004; Park et al., 2005) and is highly nutritious (Moreno et al., 2002; Kawashima and Soares, 2003; Di Noia, 2014).

Interactions between nutrients may appear at different physiological levels. The uptake of nutrients from the soil solution by the roots represents the first level of possible interaction. Mineral nutrients are usually taken up in the form of soluble salts, i.e., as cations or anions. Differences in charge leads to antagonisms and synergisms between ions and many nutrient interactions may be driven by a balance of charge. An increased uptake of an anion may lead to the decrease of nutrients taken up as cations or an increase of another anion and *vice versa*. Additionally, ion transporters usually do not exclusively transport one single nutrient as their substrate but also translocate others with similar molecular structure, though usually with a lower affinity. Therefore the deficiency or complete absence of the preferred ion might lead to the transport and subsequent accumulation of another ion that would usually be outcompeted as a substrate. This is true, for example for sulfate transporters in the plasma membrane of roots, which have been shown to also transport selenate and molybdate (Shibagaki et al., 2002; Shinmachi et al., 2010). Due to its similar size, selenium may replace sulfur in many molecules (White et al., 2004).

Studies on nutrient interactions almost exclusively apply alterations of the rhizospheric concentration of the nutrient of interest to study the impact on uptake and metabolism of other nutrients. This always bears the risk of observing indirect effects due to changes in growth. Studies on sulfur offer the possibility of supplying plants with sulfur gases as an additional or sole source of reduced sulfur. In the present study H_2S was used in concentrations that were below toxic levels but high enough to cover the bulk requirement of the plant for organic sulfur (also see Maas et al., 1987; Westerman et al., 2000). If sulfur is present in sufficient concentrations in the root medium, H_2S fumigation typically leads to a partial down-regulation of sulfate uptake by the roots. However, if sulfur is absent in the root medium, H_2S can serve as a source for sulfur and enable normal growth. In many industrial regions in the world significant amounts of H_2S are present in the atmosphere and might have an impact on the nutritional quality of crops.

The aim of the present study was to examine the separate and interactive effects of rhizospheric and atmospheric sulfur nutrition on the tissue content and shoot-to-root partitioning of other essential macro- and micronutrients. The results will help to distinguish between nutrients that are directly affected by the presence or absence of sulfate and nutrients, which are coupled to the changes in growth and organic sulfur caused by different sulfur supply.

MATERIALS AND METHODS

Plant Material, Growth Conditions, and Growth Analysis

Brassica pekinensis (Lour.) Rupr. cv. Kasumi F1 (Nickerson-Zwaan, Made, The Netherlands) was germinated in vermiculite. Ten day-old seedlings were grown in a 25% Hoagland nutrient solution (pH 5.9), consisting of 1.25 mM $Ca(NO)_3.4H_2O$, 1.25 mM KNO_3, 0.25 mM KH_2PO_4, 11.6 μM H_3BO_3, 2.4 μM $MnCl_2.4H_2O$, 0.24 μM $ZnSO_4.7H_2O$, 0.08 μM $CuSO_4.5H_2O$, 0.13 μM $Na_2MoO_4.2H_2O$, and 22.5 μM Fe^{3+}-EDTA containing either 0.5 mM (+S) or 0 mM (-S) $MgSO_4.7H_2O$. Plants were grown in 13 l containers (10 sets of plants per container, three plants per set) in climate-controlled fumigation cabinets for 11 days and fumigated with 0 or 0.2 μl l^{-1} H_2S. Day/night temperatures were 21/18°C, relative humidity was 60–70% and the photoperiod was 14 h at a photon fluence rate of 300 ± 20 μmol m^{-2} s^{-1} (within the 400–700 nm range) at plant height, supplied by Philips HPI-T (400 W) lamps.

For determination of the dry matter content fresh plant tissue was dried at 80°C for 24 h and stored in a desiccator for further use.

Analysis of Mineral Nutrient Content

Dried plant tissue (0.2–0.5 g) was digested with 5 ml of nitric acid/perchloric acid (87:13, v/v; 70% concentration, trace analysis grade; Fisher Scientific; Zhao et al., 1994). The digest solution samples were analyzed for mineral nutrients by inductively coupled plasma mass spectrometry (ICP-MS) and inductively coupled plasma atomic emission spectrometry (ICP-AES) analysis. Repeat samples were carried out every 10 samples; blanks and standard reference material (NIST 1567, a wheat flour) were used for quality control.

Inductively coupled plasma analysis was carried out using a 7500ce Octopole Reaction System ICP-MS apparatus (Agilent Technologies). The sample introduction system consisted of a micromist glass concentric nebulizer, quartz Scott-type double-pass spray chamber at 2°C, and nickel sample (1 mm) and skimmer (0.4 mm cones). Operating parameters were optimized daily using a tune solution containing 1 μg l^{-1} cerium, lithium, tellurium, and yttrium. Other instrument conditions were radiofrequency forward power of 1550, sample depth of 8.0 mm, carrier gas flow rate of 0.89 l min^{-1}, reaction gas flow rate (H_2) of 4 ml min^{-1} or (helium) of 4.5 ml min^{-1}. An internal standard (500 μg l^{-1} germanium) was used to correct for signal drift. The analytical procedures gave satisfactory values for the standard reference materials.

Mineral nutrient contents were measured from dried material. These contents were multiplied with the average dry matter content to calculate the contents based on fresh weight.

Statistical Analysis

One-way-analysis of variance (ANOVA) was used to test for significant differences in growth parameters (**Table 1**) and an

TABLE 1 | The effect of sulfur deprivation (−S) and H_2S fumigation on total biomass (fresh weight), shoot-to-root ratio and dry matter content (DMC) of shoot and roots of seedlings of Chinese cabbage.

	Total biomass (FW)	Shoot-to-root ratio	DMC shoot	DMC roots
Control	3.14 ± 0.28a	7.64 ± 0.75a	6.63 ± 0.11a	6.72 ± 0.10a
−S	0.58 ± 0.31b	4.25 ± 0.41b	11.93 ± 0.30b	8.11 ± 0.59b
H_2S	3.15 ± 0.34a	7.31 ± 1.04a	6.75 ± 0.05a	6.25 ± 0.12a
−S H_2S	2.71 ± 0.42a	5.42 ± 0.99c	6.96 ± 0.86a	6.45 ± 0.14a

Data represent the mean (± SD) of nine measurements with three plants in each. Data derived from Shahbaz et al. (2014).Values with different letters are significantly different($p < 0.05$; one-way-ANOVA,Tukey's multiple comparison as post-hoc test).

TABLE 2 | The effect of sulfur deprivation and H_2S fumigation on mineral nutrient content in [μmol g dry weight^{-1}] of shoot and roots of seedlings of Chinese cabbage.

Shoot	Control	−S	H_2S	−S H_2S
Ca	799 ± 48	600 ± 16	792 ± 17	770 ± 31
Cu	0.18 ± 0.05	0.21 ± 0.02	0.12 ± 0.01	0.12 ± 0.01
Fe	1.61 ± 0.24	1.13 ± 0.07	1.52 ± 0.07	1.23 ± 0.21
K	2046 ± 137	870 ± 31	2070 ± 14	1683 ± 65
Mg	203 ± 14	142 ± 3	199 ± 7	182 ± 2
Mn	2.29 ± 0.06	2.12 ± 0.10	2.22 ± 0.15	1.94 ± 0.16
Mo	0.06 ± 0.004	0.40 ± 0.03	0.05 ± 0.002	0.21 ± 0.01
Na	9.11 ± 1.30	11.03 ± 0.38	9.10 ± 0.41	7.73 ± 0.36
P	221 ± 16	193 ± 5	232 ± 11	212 ± 1
S	279 ± 17	29 ± 2	301 ± 17	113 ± 9
Zn	0.33 ± 0.05	0.46 ± 0.02	0.30 ± 0.03	0.50 ± 0.02
Roots	**+S**	**−S**	**+S H_2S**	**−S H_2S**
Ca	487 ± 42	289 ± 15	397 ± 37	371 ± 21
Cu	0.38 ± 0.02	0.61 ± 0.01	0.44 ± 0.02	0.72 ± 0.11
Fe	17.8 ± 2.6	34.0 ± 1.6	19.9 ± 5.1	21.6 ± 2.0
K	1681 ± 5	1498 ± 59	1944 ± 164	1544 ± 44
Mg	186 ± 11	122 ± 4	202 ± 14	135 ± 2
Mn	19.7 ± 1.7	29.4 ± 1.3	20.8 ± 3.9	19.7 ± 1.6
Mo	0.07 ± 0.001	0.96 ± 0.07	0.08 ± 0.001	0.45 ± 0.02
Na	8.4 ± 0.4	11.3 ± 0.92	10.3 ± 1.1	9.5 ± 0.6
P	309 ± 4	375 ± 10	352 ± 35	316 ± 3
S	351 ± 14	68 ± 1	330 ± 29	100 ± 3
Zn	0.72 ± 0.09	1.40 ± 0.09	0.73 ± 0.07	1.04 ± 0.02

*Data represent the mean (± SD) of three measurements with three plants in each. Relative responses and significance are shown in **Table 3**.*

Unpaired Student's *t*-test to compare nutrient contents of the treatments (-S, H_2S, -S H_2S) with the control conditions (+S; **Table 3**). A two-way-ANOVA was performed to analyze the contribution of rhizosperic and atmospheric sulfur supply to the total variance in nutrient contents (**Table 4**). The changes in sulfur and potassium content were correlated using a linear regression (**Figure 2**). All analyses were performed using GraphPad Prism (GraphPad Software, San Diego, CA, USA).

RESULTS AND DISCUSSION

Understanding all interactions between plant nutrients remains a challenge (Baxter, 2009; Maathuis, 2009) but is essential to improve nutrient use efficiency of agricultural and horticultural systems (Reich et al., 2014a). A major constraint in studying nutrient-nutrient interactions is the effect of nutrient availability on plant growth. Decreasing the tissue content of an essential nutrient below a critical level will lead to growth impairment and consequently the changes in content of other nutrients can be a direct cause of the absence of this nutrient or an indirect result of the impaired growth. The regulation of sulfate uptake and sulfur metabolism is presumed to be interconnected with plant development (Hawkesford, 2012).

This study presents results obtained from an experimental set-up in which inorganic sulfur status was uncoupled from effects on growth and the organic sulfur pool. H_2S fumigation serves as a reduced sulfur source to plants and leads to a replenishing of the organic sulfur fraction in shoots and, to a lesser extent, also in the roots whilst leaving inorganic sulfur pools (*viz.* sulfate) largely unaffected (Shahbaz et al., 2014). This creates a situation in which effects of sulfate status can be disentangled from effects of growth (**Table 1**; De Kok et al., 2007). In sulfur deficiency, the increase of calcium, copper, iron, magnesium, manganese, sodium, and phosphorus in shoots on a fresh weight basis was completely reversed if plants were supplied with H_2S (**Figure 1**, **Tables 2**, **3**). On a dry weight basis all these nutrients, except copper and sodium, were actually decreased. A two-way ANOVA showed that the variation in zinc and molybdenum content was mainly caused by rhizospheric sulfur supply (**Table 4**). The large increase of molybdenum is known to be caused by the affinity of sulfate transporters for molybdate (Leggett and Epstein, 1956; Fitzpatrick et al., 2008; Shinmachi et al., 2010). The other way around, an excessive sulfur fertilization can lead to molybdenum deficiency (MacLeod et al., 1997). Interestingly, H_2S exposure counteracted the increase of molybdenum under sulfur deficiency, although it did not completely reverse it (**Figure 1**, **Tables 2**, **3**). It is a well-known phenomenon that atmospheric, sub-lethal H_2S concentrations enable plants to maintain sufficient levels of organic sulfur compounds in leaves but do not necessarily lead to a complete down-regulation of gene expression of the sulfate transporters and sulfate uptake capacity (De Kok et al., 1997; Buchner et al., 2004; Koralewska et al., 2008). The effect of H_2S on molybdenum levels in the present study could be due to this partial down-regulation of the sulfate transporters or to an effect of growth as proposed for the other nutrients. The strong effect of sulfur deficiency observed on zinc found in the present study is a less studied phenomenon. One possible explanation for this observation is a change in rhizosphere pH. It is well-known that zinc uptake negatively correlates with rhizosphere pH (Lucas and Davis, 1961; Marschner, 1993), which is the likely reason for higher zinc uptake under ammonium nutrition and phosphorus deficiency which both lead to an acidification of the rhizosphere (Alloway, 2009; Reich et al., 2016). Measurements with H^+-electrodes showed that sulfur

TABLE 3 | Relative effect of sulfur deficiency (–S), H_2S fumigation on the content of mineral nutrients in shoot and roots of seedlings of Chinese cabbage.

Element \ Treatment	Shoot (dry weight basis)			Roots (dry weight basis)			Shoot (fresh weight basis)			Roots (fresh weight basis)		
	–S	H_2S	–S H_2S	–S	H_2S	–S H_2S	–S	H_2S	–S H_2S	–S	H_2S	–S H_2S
Ca	**25****	1n.s.	1n.s.	**41****	**19***	**24***	**35*****	1n.s.	2n.s.	**28****	**24***	**27****
Cu	16n.s.	35n.s.	36n.s.	**62*****	**18****	**92****	**109****	34n.s.	32n.s.	**95*****	10n.s.	**84****
Fe	**30***	6n.s.	24n.s.	**91*****	26n.s.	21n.s.	26n.s.	4n.s.	20n.s.	**130*****	4n.s.	16n.s.
K	**57*****	1n.s.	**18***	**11****	**16***	**8****	**23****	3n.s.	**14***	**8***	8n.s.	**12****
Mg	**30****	2n.s.	10n.s.	**34*****	9n.s.	**27****	**26****	0	6n.s.	**21****	1n.s.	**30*****
Mn	7n.s.	3n.s.	**15***	**50****	5n.s.	0	**67*****	1n.s.	11n.s.	**81*****	2n.s.	4n.s.
Mo	**575*****	**19***	**250*****	**1300*****	**23*****	**553*****	**1112*****	17n.s.	**267*****	**1589*****	14n.s.	**527*****
Na	21n.s.	0	15n.s.	**34****	**22***	13n.s.	**118*****	2n.s.	11n.s.	**62****	13n.s.	9n.s.
P	**12***	4n.s.	4n.s.	**22*****	14n.s.	2n.s.	**58*****	7n.s.	1n.s.	**47*****	6n.s.	2n.s.
S	**90*****	8n.s.	**59*****	**81*****	6n.s.	**72*****	**81*****	10n.s.	**57*****	**77*****	12n.s.	**73*****
Zn	**40***	9n.s.	**53****	**94*****	1n.s.	**45****	**150*****	7n.s.	**61****	**134*****	6n.s.	**39****

*Data expressed as relative change in % to control levels. Relative increase compared to the control is accentuated in orange, relative decrease in blue. Significant difference from the control is indicated by bold font and coloration (Unpaired Student's t-test on original values; *p < 0.05, **p < 0.01, ***p < 0.001).*

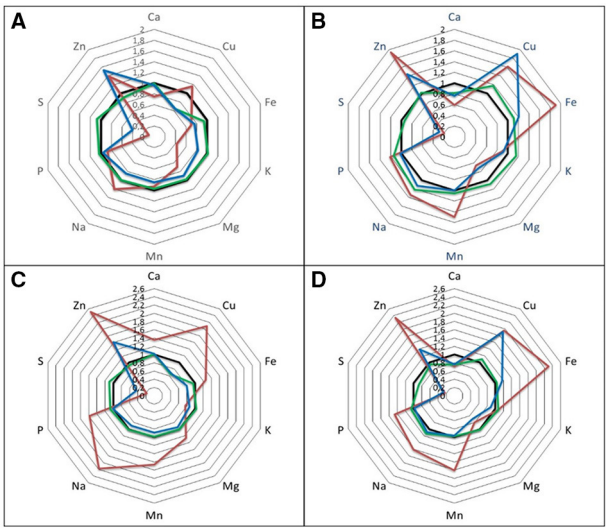


FIGURE 1 | The effect of sulfur deprivation and H_2S fumigation on mineral nutrient content in shoot and roots of seedlings of Chinese cabbage. Radar diagrams showing response ratios relative to control conditions. Shoot **(A,C)**; roots **(B,D)**; dry weight basis **(A,B)**; fresh weight basis **(C,D)**. Control (black); H_2S (green); –S (red); –S + H_2S (blue). Molybdenum was excluded from this figure due to its extraordinary large changes. For absolute contents see **Table 2**.

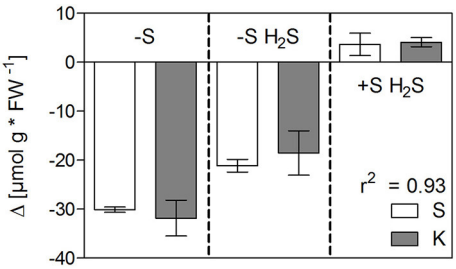


FIGURE 2 | Strong positive correlations between changes in sulfur and potassium contents in response to sulfur deficiency and H_2S fumigation in the shoot of seedlings of Chinese cabbage. The values for sulfur are multiplied by two to account for the divalency of sulfate. Data represent the mean (±SD) change in content relative to control conditions of three measurements with three plants in each. The goodness of fit (r^2) of a linear regression is indicated.

deficiency also leads to a lower pH at the roots of *B. pekinensis* seedlings (Reich et al., 2014b), which could increase zinc uptake. Another possibility for the increased content of transition metals under sulfur deficiency, and a prevention of such by H_2S, is their reactivity with and mutual detoxification by reduced sulfur compounds. Particularly cysteine-rich polypeptides possessing sulfhydryl groups (-SH) are highly reactive with transition metals (Steffens, 1990). Under sulfur deficiency these compounds are less abundant while H_2S fumigation usually leads to a restock or even higher concentrations (Buchner et al., 2004). This might explain why copper levels are recovered by H_2S fumigation (**Figure 1**, **Table 2**) but not why zinc levels are still higher under sulfur deficiency and H_2S fumigation. Both, molybdenum and zinc, are co-factors of important enzymes and zinc is involved in auxin biosynthesis (Mendel and Hänsch, 2002; Broadley et al., 2007). The metabolic consequences of an increase of these micronutrients under sulfur deficiency should be further investigated.

The only nutrient besides sulfur itself that significantly decreased in concentration in shoots under sulfur deficiency on both fresh and dry weight basis was potassium and this decrease was only partly reversed by H_2S fumigation. Interestingly, potassium decreased to about the same extent as sulfur, if its content was multiplied by two in order to take the divalency of sulfate into account (**Figure 2**). Additionally, variation in potassium content in shoots was mainly caused by rhizospheric sulfur supply (**Table 4**). Potassium therefore seemed to compensate for the changes in sulfate and to play

TABLE 4 | Results of a two-way-ANOVA showing the contribution of rhizopsheric (R) and atmospheric (A) sulfur supply and their interaction (I) to the total variance in mineral nutrient content in shoot and roots of seedlings of Chinese cabbage (%; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; Bonferroni's multiple comparison as *post-hoc* test).

	Dry weight basis						Fresh weight basis					
	Shoot			Roots			Shoot			Roots		
	R	***A***	***I***	***R***	***A***	***I***	***R***	***A***	***I***	***R***	***A***	***I***
Ca	42	23	27	56	0	33	36	28	31	38	20	29
p-value	***	***	**	***	n.s.	**	***	***	***	**	*	**
F-value	37.9	20.5	24.2	39.6	0.1	23.3	59.3	46.8	52.3	18.8	10.0	12.8
Cu	2	73	3	80	10	1	21	53	20	92	0	1
p-value	n.s.	***	n.s.	***	*	n.s.	***	***	***	***	n.s.	n.s.
F-value	0.9	26.2	0.9	64.7	7.9	0.5	31.5	78.7	29.2	109.1	0.0	1.7
Fe	65	0	4	43	14	28	2	45	31	41	25	28
p-value	**	n.s.	n.s.	**	*	**	n.s.	**	**	***	***	***
F-value	16.7	0.0	1.0	24.4	8.2	15.9	0.7	16.2	11.3	53.2	31.6	36.0
K	64	18	16	57	17	8	81	8	2	11	11	56
p-value	***	***	***	***	*	n.s.	***	*	n.s.	n.s.	n.s.	**
F value	304.0	87.1	77.4	31.4	8.9	4.4	75.9	7.7	2.2	3.9	3.9	20.4
Mg	61	13	19	91	4	0	16	39	39	86	2	4
p-value	***	**	**	***	*	n.s.	**	***	***	***	n.s.	n.s.
F-value	71.7	15.4	22.2	152.1	7.5	0.1	17.8	44.3	44.2	88.6	2.4	3.8
Mn	46	13	3	24	23	36	20	39	37	28	34	31
p-value	*	n.s.	n.s.	*	*	**	***	***	***	***	***	***
F-value	9.6	2.8	0.7	10.6	10.2	16.0	56.4	108.8	103.3	35.4	42.4	38.7
Mo	76	13	10	75	12	13	57	22	20	66	16	17
p-value	***	***	***	***	***	***	***	***	***	***	***	***
F-value	727.8	120.9	96.2	842.2	131.6	148.7	737.5	281.0	259.2	777.7	193.3	203.9
Na	1	40	39	19	0	53	25	36	38	32	15	43
p-value	n.s.	**	**	*	n.s.	**	***	***	***	**	**	***
F-value	0.4	15.4	15.2	5.4	0.0	15.2	110.9	162.0	170.6	25.3	12.3	34.3
P	52	20	3	6	2	69	28	26	43	23	27	45
p-value	**	*	n.s.	n.s.	n.s.	**	***	***	***	***	***	***
F-value	16.6	6.4	1.1	2.1	0.6	24.1	85.4	79.4	129.2	35.4	42.4	70.5
S	92	5	2	98	0	1	93	5	1	98	0	1
p-value	***	***	**	***	n.s.	*	***	***	*	***	n.s.	*
F-value	862.5	51.6	17.8	743.4	0.4	7.7	739.4	39.0	6.9	729.6	2.8	10.3
Zn	87	0	4	76	9	11	73	14	11	62	20	16
p-value	***	n.s.	n.s.	***	**	**	***	***	***	***	***	***
F-value	75.8	0.2	3.6	131.6	16.2	18.3	287.9	56.6	41.6	198.1	63.6	50.3

the role of a counter-ion. This is supported by studies on isolated vacuoles (Kaiser et al., 1989). Sulfate application also increased potassium levels in e.g., alfalfa (Razmjoo and Henderlong, 1997). In studies with Norway spruce potassium, magnesium, and manganese were increased in the endodermis and mesophyll cells if sulfate was increased due to SO_2 fumigation. This led the authors to postulate that those cations act as counter-ions for sulfate accumulation in vacuoles (Slovik et al., 1996; Bäucker et al., 2003). While potassium was decreased together with sulfur in shoots under sulfur deficiency, calcium was increased (**Figure 1**, **Table 3**). Potassium, calcium, and magnesium are known to behave antagonistically in many cases due to their common positive charge, especially at the level of uptake (Jakobsen, 1993; Marschner, 2012). Indeed, contents of the single cations in the shoot differed proportionally more between sulfur sufficient and deprived conditions, than their sum (**Figure 3**).

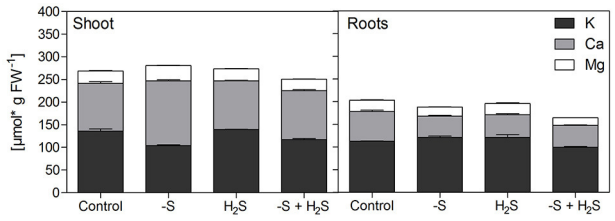


FIGURE 3 | The effect of sulfur deficiency and H_2S fumigation on the cation balance in shoot and roots of seedlings of Chinese cabbage. Data represent the mean (±SD) of the contents of potassium (K), calcium (Ca), and magnesium (Mg) in shoot (left) and root material (right) of three measurements with three plants in each. The contents of Ca and Mg were doubled to account for their divalency.

The changes in growth due to the manipulation of rhizospheric and atmospheric sulfur supply seem to be the main driver of the content of most nutrients with the few above-mentioned exceptions which are linked to clear physiochemical mechanisms. However, looking at the changes in root tissue (**Figure 1**, **Table 3**) indicates a more complex picture. The relationship discussed above between sulfur and potassium was not found here. Instead, a relative accumulation of potassium was observed while calcium and magnesium decreased. However, the sum of decrease of the two cations was not enough to compensate for the decrease in sulfur, as it was for potassium in shoots. We assume that the xylem loading of potassium and its translocation to the shoot are partly determined by the amount of sulfate translocated and accumulating in the shoot, while potassium uptake by the roots is not. Therefore, potassium accumulates in the roots when sulfate is absent.

Some authors proposed a cross-talk between sulfur and iron uptake and metabolism (Forieri et al., 2013) due to the cooperative role of both nutrients in plant metabolism, for example in iron-sulfur clusters of proteins in the electron transport chain. In the present study, however, no significant changes in iron in leaves were observed (on a fresh weight basis) under sulfur deficiency, while sulfur decreased at 81%. On the contrary, in roots a 130% increase in iron was observed under sulfur deficiency, possibly indicating a declined sink strength of the shoot for iron due to the lack of sulfur. H_2S completely reversed this effect. Also Zuchi et al. (2015) observed a decrease of iron in plants subjected to sulfur deficiency, however, here the content of the nutrients was expressed on a plant basis. As usual, sulfur deficiency led to severe impairment of growth in that study and a lower nutrient content on a plant level is to be expected. Dividing the iron content by the dry plant biomass given by Zuchi et al. (2015) revealed that that iron content was indeed also decreased by 55% on a dry weight basis, which was comparable to the decrease of 30% observed in the current study (**Table 2**). This decrease, however, disappeared if the large increase in dry matter content upon sulfur deficiency (**Table 1**) was taken into account and the content was calculated on a fresh weight basis, which presents a better estimation of the actual concentration (**Table 2**).

While increases in manganese, sodium, phosphorus, and zinc due to sulfur deficiency were prevented by H_2S fumigation in shoot and roots, the increased copper levels in roots remained completely unaffected. Both copper and zinc increase the uptake of sulfate (Shahbaz et al., 2010; Stuiver et al., 2014) and, as the present study shows, sulfate deprivation in turn led to an increase in concentration of these transition metals in root and shoot tissues under non-toxic concentrations of these micronutrients in the growing medium. Exposure with H_2S partly ameliorated this effect of sulfur deficiency, however, differently for copper and zinc. In response to H_2S exposure, sulfur deficient plants showed copper levels in the shoot similar to that of sulfur sufficient plants. In the roots however, the increased copper levels were still maintained. The different interactions of zinc and copper with sulfate uptake and assimilation need further clarification.

CONCLUSIONS

Sulfur deficiency has a diverse impact on the whole ionome of *B. pekinensis* with important implications for yield quality. By combining atmospheric and rhizospheric sulfur supply we were able to distinguish between nutrients on the basis of their direct or indirect interaction with the presence of sulfate. H_2S fumigation with simultaneous sulfate deprivation revealed that most nutrients change due to growth impairment and changes in dry matter content under sulfur deficiency, rather than a direct interaction with sulfate. Potassium was the only nutrient that was decreased together with total sulfur under sulfur deficiency and showed a strong positive correlation with sulfur content. As sulfate represents the bulk part of total sulfur, these results suggests that potassium acts as the main counter-ion for the characteristically high sulfate levels in leaves of *Brassica*. Besides molybdenum also zinc, a crucial nutrient for human nutrition, was strongly increased by sulfur deficiency independent of changes in growth. Lower root surface pH under sulfur deficiency and a lower abundance of organic sulfur compounds, which could react with zinc are possible mechanisms.

AUTHOR CONTRIBUTIONS

The experimental set up was designed by LD, MR, MS, and MH. The data for this study was acquired by MS and SP, analyzed by MR and interpreted by MR, LD, and MH. The manuscript was written by MR and LD with input from all other authors. DP practically supported the work. All authors gave their final approval for publication as well as agree to be accountable for the accuracy and integrity of the work.

ACKNOWLEDGMENTS

MR is supported by the Marie Sklodowska Curie Initial Training Network BIONUT. Work at Rothamsted Research is supported via the 20:20 Wheat® Programme by the UK Biotechnology and Biological Sciences Research Council.

REFERENCES

Alloway, B. J. (2009). Soil factors associated with zinc deficiency in crops and humans. *Environ. Geochem. Health* 31, 537–548. doi: 10.1007/s10653-009-9255-4

Bäucker, E., Köhler, R., Matschullat, J., Wienhaus, O., and Zimmermann, F. (2003). Sulphate and cation accumulation in mesophyll and endodermis vacuoles of spruce needles–effects on nutrient supply after SO_2 fumigation. *Trees* 17, 1–10. doi: 10.1007/s00468-002-0199-x

Baxter, I. (2009). Ionomics: studying the social network of mineral nutrients. *Curr. Opin. Plant Biol.* 12, 381–386. doi: 10.1016/j.pbi.2009.05.002

Bogdanov, S. (1899). On the sulphur in plants. *Exp. Sta. Rec.* 11, 723–724.

Broadley, M. R., White, P. J., Hammond, J. P., Zelko, I., and Lux, A. (2007). Zinc in plants. *New Phytol.* 173, 677–702. doi: 10.1111/j.1469-8137.2007.01996.x

Buchner, P., Stuiver, C. E. E., Westerman, S., Wirtz, M., Hell, R., Hawkesford, M. J., et al. (2004). Regulation of sulfate uptake and expression of sulfate transporter genes in *Brassica oleracea* as affected by atmospheric H_2S and pedospheric sulfate nutrition. *Plant Physiol.* 136, 3396–3408. doi: 10.1104/pp.104.046441

De Kok, L. J., Castro, A., Durenkamp, M., Koralewska, A., Posthumus, F. S., Stuiver, C. E. E., et al. (2005). Pathways of plant sulfur uptake and metabolism-an overview. *Landbauforsch. Volk.* 28, 5–13.

De Kok, L. J., Durenkamp, M., Yang, L., and Stulen, I. (2007). "Atmospheric sulfur," in *Sulfur in Plants. An Ecological Perspective,* eds M. J. Hawkesford and L. J. De Kok (Dordrecht: Springer), 91–106.

De Kok, L. J., Stuiver, C. E. E., Rubinigg, M., Westerman, S., and Grill, D. (1997). Impact of atmospheric sulfur deposition on sulfur metabolism in plants: H_2S as sulfur source for sulfur deprived *Brassica oleracea* L. *Bot. Acta* 110, 411–419. doi: 10.1111/j.1438-8677.1997.tb00657.x

Di Noia, J. (2014). Defining powerhouse fruits and vegetables: a nutrient density approach. *Prev. Chronic Dis.* 11:130390. doi: 10.5888/pcd11.130390

Fitzpatrick, K. L., Tyerman, S. D., and Kaiser, B. N. (2008). Molybdate transport through the plant sulfate transporter SHST1. *FEBS Lett.* 582, 1508–1513. doi: 10.1016/j.febslet.2008.03.045

Forieri, I., Wirtz, M., and Hell, R. (2013). Toward new perspectives on the interaction of iron and sulfur metabolism in plants. *Front. Plant Sci.* 4:357. doi: 10.3389/fpls.2013.00357

Hart, E. B., and Peterson, W. H. (1911). The sulfur requirements of farm crops in relation to the soil and air supply. *J. Am. Chem. Soc.* 33, 49–564. doi: 10.1021/ja02217a015

Hawkesford, M. J. (2012). "Sulfate uptake and assimilation–whole plant regulation," in *Sulfur Metabolism in Plants*, eds L. J. De Kok, M. Tausz, M. J. Hawkesford, R. Hoefgen, M. T. McManus, R. Norton, H. Rennenberg, K. Saito, E. Schnug, and L. Tabe (Dordrecht: Springer), 11–24.

Hawkesford, M. J., and De Kok, L. J. (2006). Managing sulphur metabolism in plants. *Plant Cell Environ.* 29, 382–395. doi: 10.1111/j.1365-3040.2005.01470.x

Hell, R. (1997). Molecular physiology of plant sulfur metabolism. *Planta* 202, 138–148. doi: 10.1007/s004250050112

Honsel, A., Kojima, M., Haas, R., Frank, W., Sakakibara, H., Herschbach, et al. (2012). Sulphur limitation and early sulphur deficiency responses in poplar: significance of gene expression, metabolites, and plant hormones. *J. Exp. Bot.* 63, 1873–1893. doi: 10.1093/jxb/err365

Jakobsen, S. T. (1993). Interaction between plant nutrients: III. Antagonism between potassium, magnesium and calcium. *Acta Agric. Scand. B Soil Plant Sci.* 43, 1–5. doi: 10.1080/09064719309410223

Kaiser, G., Martinoia, E., Schröppel-Meier, G., and Heber, U. (1989). Active transport of sulfate into the vacuole of plant cells provides halotolerance and can detoxify SO_2. *J. Plant Physiol.* 133, 756–763. doi: 10.1016/S0176-1617(89)80085-9

Kawashima, L. M., and Soares, L. M. V. (2003). Mineral profile of raw and cooked leafy vegetables consumed in Southern Brazil. *J. Food Compost. Anal.* 16, 605–611. doi: 10.1016/S0889-1575(03)00057-7

Kopriva, S. (2006). Regulation of sulfate assimilation in Arabidopsis and beyond. *Ann. Bot.* 97, 479–495. doi: 10.1093/aob/mcl006

Kopriva, S., Suter, M., von Ballmoos, P., Hesse, H., Krähenbühl, U., Rennenberg, H., et al. (2002). Interaction of sulfate assimilation with carbon and nitrogen metabolism in *Lemna minor*. *Plant Physiol.* 130, 1406–1413. doi: 10.1104/pp.007773

Koralewska, A., Stuiver, C. E. E., Posthumus, F. S., Kopriva, S., Hawkesford, M. J., and De Kok, L. J. (2008). Regulation of sulfate uptake, expression of the sulfate transporters Sultr1; 1 and Sultr1; 2, and APS reductase in Chinese cabbage (*Brassica pekinensis*) as affected by atmospheric H_2S nutrition and sulfate deprivation. *Funct. Plant. Biol.* 35, 318–327. doi: 10.1071/FP07283

Leggett, J. E., and Epstein, E. (1956). Kinetics of sulfate absorption by barley roots. *Plant Physiol.* 31, 222–226.

Leustek, T., Martin, M. N., Bick, J. A., and Davies, J. P. (2000). Pathways and regulation of sulfur metabolism revealed through molecular and genetic studies. *Annu. Rev. Plant Biol.* 51, 141–165. doi: 10.1146/annurev.arplant.51.1.141

Lucas, R. E., and Davis, J. (1961). Relationships between pH values of organic soils and availabilities of 12 plant nutrients. *Soil Sci.* 92, 177–182. doi: 10.1097/00010694-196109000-00005

Maas, F. M., De Kok, L. J., Peters, J. L., and Kuiper, P. J. (1987). A comparative study on the effects of H_2S and SO_2 fumigation on the growth and accumulation of sulphate and sulphydryl compounds in *Trifolium pratense* L., *Glycine max* Merr. and *Phaseolus vulgaris* L. *J. Exp. Bot.* 38, 1459–1469. doi: 10.1093/jxb/38.9.1459

Maathuis, F. J. (2009). Physiological functions of mineral macronutrients. *Curr. Opin. Plant Biol.* 12, 250–258. doi: 10.1016/j.pbi.2009.04.003

MacLeod, J. A., Gupta, U. C., and Stanfield, B. (1997). "Molybdenum and sulfur relationships in plants," in *Molybdenum in Agriculture,* ed U. C. Gupta (Cambridge: Press Syndicate of the University of Cambridge), 229–244.

Marschner, H. (1993). "Zinc uptake from soils," in *Zinc in Soils and Plants,* ed A. D. Robson (Dordrecht: Springer), 59–77.

Marschner, H. (2012). *Marschner's Mineral Nutrition of Higher Plants*, Vol. 89, ed P. Marschner (London: Academic Press).

Mendel, R. R., and Hänsch, R. (2002). Molybdoenzymes and molybdenum cofactor in plants. *J. Exp. Bot.* 53, 1689–1698. doi: 10.1093/jxb/erf038

Moreno, D. A., Víllora, G., Hernández, J., Castilla, N., and Monreal, L. R. (2002). Yield and chemical composition of Chinese cabbage in relation to thermal regime as influenced by row covers. *J. Am. Soc. Hortic. Sci.* 127, 343–348.

Park, B. J., Liu, Z., Kanno, A., and Kameya, T. (2005). Genetic improvement of Chinese cabbage for salt and drought tolerance by constitutive expression of a *B. napus* LEA gene. *Plant Sci.* 169, 553–558. doi: 10.1016/j.plantsci.2005.05.008

Rakow, G. (2004). "Species origin and economic importance of *Brassica*," in *Brassica*, eds E. G. Pua and C. J. Douglas (Berlin; Heidelberg: Springer), 3–11.

Razmjoo, K., and Henderlong, P. R. (1997). Effect of potassium, sulfur, boron, and molybdenum fertilization on alfalfa production and herbage macronutrient contents. *J. Plant Nutr.* 20, 1681–1696. doi: 10.1080/01904169709365367

Reich, M., Aghajanzadeh, T., and De Kok, L. J. (2014a). "Physiological basis of plant nutrient use efficiency – concepts, opportunities and challenges for its improvement," in *Nutrient Use Efficiency in Plants*, eds M. J. Hawkesford, S. Kopriva, and L. J. De Kok (Dordrecht: Springer International Publishing), 1–27.

Reich, M., Meerakker, A. N., Parmar, S., Hawkesford, M. J., and De Kok, L. J. (2016). Temperature determines size and direction of effects of elevated CO_2 and nitrogen form on yield quantity and quality of Chinese cabbage. *Plant Biol.* 18, 63–75. doi: 10.1111/plb.12396

Reich, M., Staal, M., Elzenga, J. T. M., and De Kok, L. J. (2014b). Using H^+-selective microelectrodes to study the adaptive response of plant roots to sulfate deficiency. *J. Cultiv. Plants* 66, 106.

Shahbaz, M., Hwei Tseng, M., Stuiver, C. E. E., Koralewska, A., Posthumus, F. S., Venema, J. H., et al. (2010). Copper exposure interferes with the regulation of the uptake, distribution and metabolism of sulfate in Chinese cabbage. *J. Plant Physiol.* 167, 438–446. doi: 10.1016/j.jplph.2009.10.016

Shahbaz, M., Stuiver, C. E. E., Posthumus, F. S., Parmar, S., Hawkesford, M. J., and Kok, L. J. (2014). Copper toxicity in Chinese cabbage is not influenced by plant sulphur status, but affects sulphur metabolism−related gene expression and the suggested regulatory metabolites. *Plant Biol.* 16, 68–78. doi: 10.1111/plb.12019

Shibagaki, N., Rose, A., McDermott, J. P., Fujiwara, T., Hayashi, H., Yoneyama, T., et al. (2002). Selenate−resistant mutants of *Arabidopsis thaliana* identify Sultr1;2, a sulfate transporter required for efficient transport of sulfate into roots. *Plant J.* 29, 475–486. doi: 10.1046/j.0960-7412.2001.01232.x

Shinmachi, F., Buchner, P., Stroud, J. L., Parmar, S., Zhao, F. J., McGrath, S. P., et al. (2010). Influence of sulfur deficiency on the expression of specific sulfate transporters and the distribution of sulfur, selenium and molybdenum in wheat. *Plant Physiol.* 153, 327–336. doi: 10.1104/pp.110.153759

Slovik, S., Hüve, K., Kindermann, G., and Kaiser, W. M. (1996). SO_2—dependent cation competition and compartmentalization in Norway spruce needles. *Plant Cell Environ*. 19, 813–824. doi: 10.1111/j.1365-3040.1996.tb00418.x

Steffens, J. C. (1990). The heavy metal-binding peptides of plants. *Ann. Rev. Plant Biol.* 41, 553–575. doi: 10.1146/annurev.pp.41.060190.003005

Stuiver, C. E. E., Posthumus, F. S., Parmar, S., Shahbaz, M., Hawkesford, M. J., and De Kok, L. J. (2014). Zinc exposure has differential effects on uptake and metabolism of sulfur and nitrogen in Chinese cabbage. *J. Soil Sci. Plant Nutr.* 177, 748–757. doi: 10.1002/jpln.201300369

Takahashi, H., Kopriva, S., Giordano, M., Saito, K., and Hell, R. (2011). Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Ann. Rev. Plant Biol.* 62, 157–184. doi: 10.1146/annurev-arplant-042110-103921

Thompson, J. F. (1967). Sulfur metabolism in plants. *Ann. Rev. Plant Physiol.* 18, 59–84. doi: 10.1146/annurev.pp.18.060167.000423

Westerman, S., De Kok, L. J., Stuiver, C. E. E., and Stulen, I. (2000). Interaction between metabolism of atmospheric H_2S in the shoot and sulfate uptake by the roots of curly kale (*Brassica oleracea*). *Physiol. Plant.* 109, 443–449. doi: 10.1034/j.1399-3054.2000.100411.x

White, P. J., Bowen, H. C., Parmaguru, P., Fritz, M., Spracklen, W. P., Spiby, R. E., et al. (2004). Interactions between selenium and sulphur nutrition in *Arabidopsis thaliana*. *J. Exp. Bot.* 55, 1927–1937. doi: 10.1093/jxb/erh192

Zhao, F. J., Hawkesford, M. J., and McGrath, S. P. (1999). Sulphur assimilation and effects on yield and quality of wheat. *J. Cereal Sci.* 30, 1–17. doi: 10.1006/jcrs.1998.0241

Zhao, F., McGrath, S. P., and Crosland, A. R. (1994). Comparison of three wet digestion methods for the determination of plant sulphur by inductively coupled plasma atomic emission spectroscopy (ICP-AES). *Commun. Soil Sci. Plan.* 25, 407–418. doi: 10.1080/00103629409369047

Zuchi, S., Watanabe, M., Hubberten, H. M., Bromke, M., Osorio, S., Fernie, A. R., et al. (2015). The interplay between sulfur and iron nutrition in tomato. *Plant Physiol.* 169, 2624–2639. doi: 10.1104/pp.15.00995

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Reich, Shahbaz, Prajapati, Parmar, Hawkesford and De Kok. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 May 2016
doi: 10.3389/fpls.2016.00581

Brassica napus Genome Possesses Extraordinary High Number of *CAMTA* Genes and *CAMTA3* Contributes to PAMP Triggered Immunity and Resistance to *Sclerotinia sclerotiorum*

*Hafizur Rahman[1], You-Ping Xu[2], Xuan-Rui Zhang[1] and Xin-Zhong Cai[1]**

[1] Institute of Biotechnology, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, China, [2] Center of Analysis and Measurement, Zhejiang University, Hangzhou, China

OPEN ACCESS

Edited by:
Naser A. Anjum,
University of Aveiro, Portugal

Reviewed by:
Zhongyun Piao,
Shenyang Agricultural University, China
Chris Gehring,
King Abdullah University of Science and Technology, Saudi Arabia

****Correspondence:***
Xin-Zhong Cai
xzhcai@zju.edu.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *04 February 2016*
Accepted: *14 April 2016*
Published: *04 May 2016*

Citation:
Rahman H, Xu Y-P, Zhang X-R and Cai X-Z (2016) Brassica napus Genome Possesses Extraordinary High Number of CAMTA Genes and CAMTA3 Contributes to PAMP Triggered Immunity and Resistance to Sclerotinia sclerotiorum. Front. Plant Sci. 7:581. doi: 10.3389/fpls.2016.00581

Calmodulin-binding transcription activators *(CAMTAs)* play important roles in various plant biological processes including disease resistance and abiotic stress tolerance. Oilseed rape (*Brassica napus* L.) is one of the most important oil-producing crops worldwide. To date, compositon of *CAMTAs* in genomes of *Brassica* species and role of *CAMTAs* in resistance to the devastating necrotrophic fungal pathogen *Sclerotinia sclerotiorum* are still unknown. In this study, 18 *CAMTA* genes were identified in oilseed rape genome through bioinformatics analyses, which were inherited from the nine copies each in its progenitors *Brassica rapa* and *Brassica oleracea* and represented the highest number of *CAMTAs* in a given plant species identified so far. Gene structure, protein domain organization and phylogentic analyses showed that the oilseed rape *CAMTAs* were structurally similar and clustered into three major groups as other plant *CAMTAs*, but had expanded subgroups *CAMTA3* and *CAMTA4* genes uniquely in rosids species occurring before formation of oilseed rape. A large number of stress response-related *cis*-elements existed in the 1.5 kb promoter regions of the *BnCAMTA* genes. *BnCAMTA* genes were expressed differentially in various organs and in response to treatments with plant hormones and the toxin oxalic acid (OA) secreted by *S. sclerotiorum* as well as the pathogen inoculation. Remarkably, the expression of *BnCAMTA3A1* and *BnCAMTA3C1* was drastically induced in early phase of *S. sclerotiorum* infection, indicating their potential role in the interactions between oilseed rape and *S. sclerotiorum*. Furthermore, inoculation analyses using Arabidopsis *camta* mutants demonstrated that *Atcamta3* mutant plants exhibited significantly smaller disease lesions than wild-type and other *Atcamta* mutant plants. In addition, compared with wild-type plants, *Atcamta3* plants accumulated obviously more hydrogen peroxide in response to the PAMP chitin and exhibited much higher expression of the CGCG-box-containing genes *BAK1* and *JIN1*, which are essential to the PAMP triggered immunity (PTI) and/or plant resistance

to pathogens including *S. sclerotiorum*. Our results revealed that *CAMTA3* negatively regulated PTI probably by directly targeting *BAK1* and it also negatively regulated plant defense through suppressing JA signaling pathway probably via directly targeting *JIN1*.

Keywords: *Brassica napus*, *CAMTA*, disease resistance, PAMP triggered immunity, *Sclerotinia sclerotiorum*

INTRODUCTION

Calcium is a ubiquitous second messenger used by plants to regulate a variety of biological processes in response to a wide range of environmental and developmental stimuli (Galon et al., 2010; Reddy et al., 2011). In response to these stimuli, Ca^{2+} signals are decoded and transmitted by several types of Ca^{2+} sensor proteins including calmodulins (CaMs), calcineurin B-like proteins (CBLs), and calcium-dependent protein kinases (CDPKs/CPKs; Kudla et al., 2010; Du et al., 2011). CaM can bind to certain transcription factors such as calmodulin-binding transcription activators *(CAMTAs)*.

CAMTAs, also referred to as signal-responsive (SR) proteins, are thought to exist in all multicellular organisms (Bouché et al., 2002; Rahman et al., 2016). Taking advantage of the rapid developing of plant genome sequencing, *CAMTA* family has been identified at genome-wide level in over 40 plant species (Bouché et al., 2002; Choi et al., 2005; Koo et al., 2009; Yang et al., 2012; Shangguan et al., 2014; Wang et al., 2015; Yang et al., 2015; Yue et al., 2015; Rahman et al., 2016). Nevertheless, composition of *CAMTAs* in many economically important crop species such as *Brassica* species is still unknown.

CAMTAs contain multiple functional domains including a CG-1 DNA-binding domain, an ankyrin (ANK) repeat domain, an IQ (Isoleucine glutamine) domain, and a CaM binding (CaMB) domain that are located in turn from the N terminus to the C terminus (Bouché et al., 2002; Choi et al., 2005; Finkler et al., 2007; Rahman et al., 2016). Most of *CAMTAs* also possess a TIG (Transcription-associated immuno globulin-like) domain (Rahman et al., 2016). *CAMTAs* specifically recognize and bind to (A/C/G)CGCG(T/C/G) or (A/C)CGTGT *cis*-elements in the promoter region of target genes, thereby regulate their expression (Yang and Poovaiah, 2002; Choi et al., 2005; Du et al., 2009). The biological functions of *CAMTAs* are being revealed but mainly in Arabidopsis, rice and tomato. The functions of *CAMTAs* were dependent on their interaction with Ca^{2+}/CaM (Choi et al., 2005; Du et al., 2009). Arabidopsis *CAMTA3* negatively regulates accumulation of salicylic acid and host plant resistance to both bacterial (Du et al., 2009) and fungal pathogens (Galon et al., 2008; Nie et al., 2012) as well as nonhost resistance to bacterial pathogen *Xanthomonas oryzae* pv. *oryzae*, probably via tuning *CBP60G*, *EDS1,* and *NDR1*-mediated defense signaling and reactive oxygen species (ROS) accumulation (Rahman et al., 2016). AtCAMTA3 signaling is modulated by ubiquitination process during regulation of plant immunity (Zhang et al., 2014). Similarly, a rice *CAMTA* OsCBT-1 negatively regulates rice resistance to blast fungal pathogen and leaf blight bacterial pathogen (Koo et al., 2009). Besides, AtCAMTA3 also plays important roles in plant defense against insect herbivore, glucose metabolism and ethylene-induced senescence in Arabidopsis (Laluk et al., 2012; Qiu et al., 2012). Arabidopsis *CAMTA1, CAMTA2,* and *CAMTA3* contribute to low temperature and freezing tolerance by activation of *CBF* (C-repeat/DRE binding factor) transcription factors (Doherty et al., 2009; Kim et al., 2013). Tomato *CAMTAs* are differentially expressed during fruit development and ripening processes and in responsive to biotic and abiotic stimuli (Yang et al., 2012, 2013). Silencing of *SlSR1* and *SlSR3L* enhances resistance to bacterial and fungal pathogens while silencing of SlSR1L leads to decreased drought stress tolerance (Li et al., 2014). Collectively, these reports clearly demonstrate that *CAMTAs*, especially *CAMTA3*, are important regulators of plant resistance to biotrophic pathogens. Nevertheless, their role in plant resistance to necrotrophic pathogens remains poorly understood.

Oilseed rape (*Brassica napus* L.) is one of the most important oil crops worldwide. Despite relatively extensive studies of *CAMTAs* in several model plant species, little is known about this gene family in oilseed rape and other *Brassica* species. Only one *CAMTA* sequence has been identified in oilseed rape to date (Bouché et al., 2002). In this study, taking advantage of completion of the oilseed rape genome sequence (Chalhoub et al., 2014), we systemically identified the *CAMTA* gene family in *B. napus* genome and performed comprehensive sequence analyses as well as functional analyses in disease resistance. Our results demonstrated that oilseed rape genome contained the highest number of *CAMTAs* in a given plant species identified so far. *BnCAMTA3A1* and *BnCAMTA3C1* were likely to be the functional homologs of *AtCAMTA3* functioning in disease resistance. Furthermore, using Arabidopsis *camta* mutants, we revealed that *CAMTA3* negatively regulated chitin-triggerred immunity and plant defense to the devastating necrotrophic pathogen *Sclerotinia sclerotiorum*, probably via directly targeting *BAK1* and *JIN1*.

MATERIALS AND METHODS

Identification of *CAMTA* Proteins in *Brassica* Species

To identify *CAMTA* protein sequences in oilseed rape, the six Arabidopsis *CAMTAs* were used as query to search by BLASTP program against *B. napus* genome databases deposited in NCBI (http://www.ncbi.nlm.nih.gov/) and the GNEOSCOPE (http://www.genoscope.cns.fr/spip/). All retrieved non-redundant sequences were collected, and subjected to conserved domain analysis using the Pfam (http://pfam.sanger.ac.uk/) and NCBI-CDD (http://www.ncbi.nlm.nih.gov/cdd) databases. These sequences were compared with Arabidopsis and tomato CAMTA proteins using ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with default settings and were viewed by

GeneDoc. Those containing a CG-1 domain, an ANK repeat domain and a CaMB domain were recognized as CAMTA proteins. CAMTAs in oilseed rape were named in accordance with their phylogenetic relationship to six Arabidopsis CAMTAs. Identification of CAMTAs in *B. rapa* and *B. oleracea*, two progenitor species of *B. napus*, was performed similarly.

Gene Structure, Protein Domain, and Phylogenetic Analyses of *BnCAMTA* Genes

The gene structure was analyzed online by the Gene Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/index.php; Guo et al., 2007). A schematic diagram of protein domain structures with functional motifs was constructed using Domain Illustrator software (http://dog.biocuckoo.org/; Ren et al., 2009). The sequence logos of CaMB domain were generated using the Geneious software (v6.1.6) package (http://www.geneious.com/). Multiple sequence alignments of the full-length CAMTA proteins from representative plant species were conducted using ClustalW. The phylogenetic tree was constructed using MEGA 5.0 (Tamura et al., 2011) with maximum likelihood (ML) method and a bootstrap test was performed with 1000 replicates.

Prediction of *cis*-Acting Elements in the *BnCAMTA* Genes

To investigate *cis*-elements in the promoter sequences of the *BnCAMTA* genes, 1.5 kb sequences upstream of the initiation codon (ATG) were collected and subjected to stress response-related *cis*-acting element online prediction analysis with Signal Scan search program in the PLACE database (http://www.dna.affrc.go.jp/PLACE/signalscan.html) and the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Plant Material and Hormone Treatments

Oilseed rape plants were grown in growth room at 22–23°C with a 16/8 h day/night photoperiod. Arabidopsis plants of Col-0 and six *CAMTA* mutants (*Atcamta1-6*) were grown in a growth chamber at 20–21°C under a 15/9 h day/night photoperiod. For *BnCAMTA* gene expression analyses, leaves of 4-week-old plants were sprayed with hormones SA (1 mM) and JA (200 μM) as well as a chemical OA (1 mM), the toxin secreted by the pathogen *S. sclerotiorum*, or 0.01% ethonal (the solvent for the above chemicals) as a control, and collected at 0, 4, 12, and 24 h after treatment. In addition, various organs of oilseed rape plants including root, stem, cotyledone, and true leaves were also sampled for gene expression analysis. All samples were immediately frozen in liquid nitrogen and stored at −80°C until RNA extraction.

S. sclerotiorum Inoculation and Plant Resistance Analyses

Leaves of 4 week-old *B. napus* and Arabidopsis plants were inoculated with mycelial plugs of 3 mm diameter of *S. sclerotiorum* as described (Saand et al., 2015a). The inoculated leaves were collected at 0, 6, and 12 h post inoculation for RNA extraction and gene expression analyses. The necrosis symptoms of the inoculated leaves were investigated and the size of lesions was measured. The inoculation analysis was performed three times, each in at least 6 plants for each treatment and gene backgrounds. For the statistical analysis of the lesion size data, ANOVA (analysis of variance) analysis was performed with SPSS software (Version 19.0, IBM, USA). Significant difference between the mean values of three independent experiments was determined with Duncan's multiple range test (DMRT; $p < 0.05$).

Detection of Chitin-Triggered Hydrogen Peroxide

The hydrogen peroxide (H_2O_2) elicited by chitin (100 μg mL^{-1}, Sigma, USA) in leaf discs of *Atcamta3* mutant and wild type Col-0 plants were measured using a Microplate Luminometer (TITERTEK BERTHOLD, Germany) following previously described protocol (Saand et al., 2015a). For each experiment, 10 leaves were collected for each genotype. All experiments were conducted three times independently. The quantitative measurement data were statistically analyzed using SPSS software and represent means ± standard error.

RNA Isolation and Gene Expression Analyses

Total RNA was extracted with Trizol reagent (TAKARA, Japan) following the manufacturer's instructions. RNA was treated with DNase I (TAKARA, Japan) and reverse-transcribed into cDNA using the PrimeScript RT reagent kit (TAKARA, Japan). The obtained cDNAs were used for gene expression analyses with semiquantitative reverse transcription PCR (RT-PCR) and quantitative real time PCR (qRT-PCR). Semiquantitative RT-PCR was performed following the program: 94°C for 5 min, followed by 32 or 28 (for internal control gene) cycles of denaturation for 50 s at 94°C, annealing for 50 s at 55°C, extension for 20 s at 72°C, and a final extension for 10 min at 72°C. The obtained products were analyzed by electrophoresis on a 1.5% agarose gel and detected under ultraviolet light. The qRT-PCR was conducted in StepOne Real-Time PCR System (Applied Biosystems, USA) using SYBER Premix Ex Taq reagents (TaKaRa, Japan) following the program: 95°C for 30 s, 95°C for 5 s, and 60°C for 45 s for 40 cycles. To normalize the sample variance, *B. napus β -Tubulin* and Arabidopsis *ACTIN8* genes served as internal controls. Relative gene expression values were calculated using the $2^{-\Delta\Delta Ct}$ method. The primers used for gene expression analyses are listed at Table S1. For the statistical analysis of the gene expression data, ANOVA analysis was performed with SPSS software (Version 19.0, IBM, USA). Significant difference between mean values was determined with DMRT ($p < 0.05$).

RESULTS

Identification of *CAMTA* Genes in *B. napus* and Its Two Progenitor Species

To identify *CAMTA* genes in *B. napus*, the six Arabidopsis CAMTAs were used as query to BLASTP search in the complete genome of *B. napus*. Based on domain composition analyses for the retrieved candidate sequences, a total of 18 CAMTA sequences were identified in *B. napus* genome, representing the

highest number of CAMTAs in a given plant species identified so far. They were named in accordance with their phylogenetic relationship with the six Arabidopsis CAMTAs and the location in subgenomes (A or C). The comprehensive information of *BnCAMTA* genes, including locus ID, gene location, length and intron number, predicted protein size, molecular weight, and isoelectric point (pI), is listed in **Table 1**. The length of the *BnCAMTA* gene sequences was 4.6–6.0 kb with three exceptions *BnCAMTA3C2* (10.0 kb), *BnCAMTA4A2* (8.9 kb), and *BnCAMTA6A* (6.7 kb), which contained significantly longer genomic sequence due to possessing an extraordinarily large intron (**Table 1**; **Figure 1**). The size of predicted BnCAMTA proteins was 919–1034 amino acids (aa) except BnCAMTAs 6A and 6C with 853 aa and BnCAMTA4A2 with 1258 aa (**Table 1**). BnCAMTA4A2 was larger due to carrying an extra N terminal sequence (**Figure 1A**). BnCAMTA proteins varied obviously in their pI value. The majority of them (12 out of 18) owned a pI of lower than 6.2, two of them (BnCAMTAs 6A and 6C) had a pI of near 7.0, while the remaining four (BnCAMTAs 4C2, 4A2, 5A, and 5C) possessed a pI of higher than 7.4 (**Table 1**), implying that while most of the BnCAMTA proteins are acidic, some of them are neutral or basic. Collectively, these results indicated that *B. napus* genome possesses much more *CAMTA* genes than other plant species, and their physico-chemical characteristics were generally conserved but with obvious exceptions.

To better understand the composition of CAMTAs in the tetraploid *B. napus*, CAMTAs in its two progenitor species *B. rapa* and *B. oleracea* was also identified using similar approaches. The results showed that the two *Brassica* species exhibited similar *CAMTA* composition, both containing 9 CAMTAs (Table S1). Comparison analysis indicated that *B. napus* genome possessed exactly the total copies of CAMTAs in its two progenitor species.

Chromosomal Location of *BnCAMTA* Gene Family

The 18 *BnCAMTA* genes were mapped on 14 oilseed rape chromosomes (Figure S1). Among them, eight were scattered each on one chromosome (A02, C02_random, A04, C04, A05, C05, A08, and A010), while the remaining 10 were distributed in five chromosomes (C06, A07, C08, A09, and C09) with each two genes in one chromosome. BnCAMTAs 4C1 and 4C2 as well as BnCMTAs 4A1 and 4A2 were located nearly each other on Chromosomes C06 and A07, respectively, while BnCMTAs 2A and 3A2 as well as BnCAMTAs 1C and 2C were distributed distantly on the two ends of Chromosomes A09 and C09, respectively (Figure S1). In addition, BnCAMTAs 3C2 and 5C lay in Chromosome C08 although the precise position of BnCAMTA5C in this chromosome remained unclear (tentatively called ChrC08_random). This result suggested that gene duplication and recombination occurred, most obviously for *BnCAMTA4s*, and contributed to *CAMTA* gene expansion in *B. napus*.

Conserved Domain and Gene Structural Analyses of BnCAMTAs

The CAMTA proteins consist of multiple predicted functional domains, evolutionally conserved in amino acid sequence and organization order. The domain structure analyses revealed that all the 18 BnCAMTA proteins contained a CG-1 DNA-binding domain in the N-terminal portion, an ankyrin repeat (ANK) domain in the middle, one or two IQ motifs and a calmodulin binding (CaMB) domain in the C-terminal region (**Figure 1A**). In addition, 10 BnCAMTAs belonging to subgroups 4, 5, and 6 contained a TIG domain, located between the N-terminal CG-1 domain and the ANK domain (**Figure 1A**). All BnCAMTA

TABLE 1 | *CAMTA* gene family in oilseed rape.

Gene name	Locus ID	Gene location	Gene length (bp)	No. of introns	No. of amino acid (aa)	Mol.Wt. (kDa)	pI
Bn*CAMTA1A*	BnaA10g22560D	ChrA10: 15173521-15178533	5013	12	1007	113.5	5.88
Bn*CAMTA1C*	BnaC09g47120D	ChrC09: 46651396-46656382	4987	13	999	112.4	5.96
Bn*CAMTA*2A	BnaA09g06760D	ChrA09: 3297678-3303234	5557	13	1019	114	6.13
Bn*CAMTA*2C	BnaC09g06280D	ChrC09: 3766467-3772427	5961	12	987	110.2	5.99
Bn*CAMTA*3A1	BnaA04g12770D	ChrA04: 10768124-10773453	5330	14	990	111	5.49
Bn*CAMTA*3C1	BnaC04g34700D	ChrC04: 36202016-36207150	5135	11	1028	115	5.48
Bn*CAMTA*3A2	BnaA09g42730D	ChrA09: 29720801-29725508	4708	11	924	104.4	5.81
Bn*CAMTA*3C2	BnaC08g35210D	ChrC08: 33166518-33176534	10020	9	947	106.6	5.73
Bn*CAMTA*4A1	BnaA07g25100D	ChrA07: 18753935-18759298	5364	13	965	107.7	5.93
Bn*CAMTA*4C1	BnaC06g26850D	ChrC06: 28406024-28411276	5253	12	976	108.5	5.86
Bn*CAMTA*4A2	BnaA07g26320D	ChrA07: 19369075-19378014	8940	21	1258	139.5	7.94
Bn*CAMTA*4C2	BnaC06g28390D	ChrC06: 29594775-29600235	5461	14	1025	115.4	7.9
Bn*CAMTA*4A3	BnaA02g13050D	ChrA02: 7109254-7114694	5441	13	1034	115.3	5.42
Bn*CAMTA*4C3	BnaC02g45620D	ChrC02_random: 1363612-1369466	5855	13	1025	114.5	5.63
Bn*CAMTA*5A	BnaA08g06080D	ChrA08: 6008140-6012801	4662	12	919	104.4	7.59
Bn*CAMTA*5C	BnaC08g47670D	ChrC08_random: 2585648-2590284	4637	12	919	104.5	7.45
Bn*CAMTA*6A	BnaA05g22990D	ChrA05: 17452572-17459296	6725	11	853	97	6.92
Bn*CAMTA*6C	BnaC05g36450D	ChrC05: 35682877-35687702	4826	11	853	97	6.84

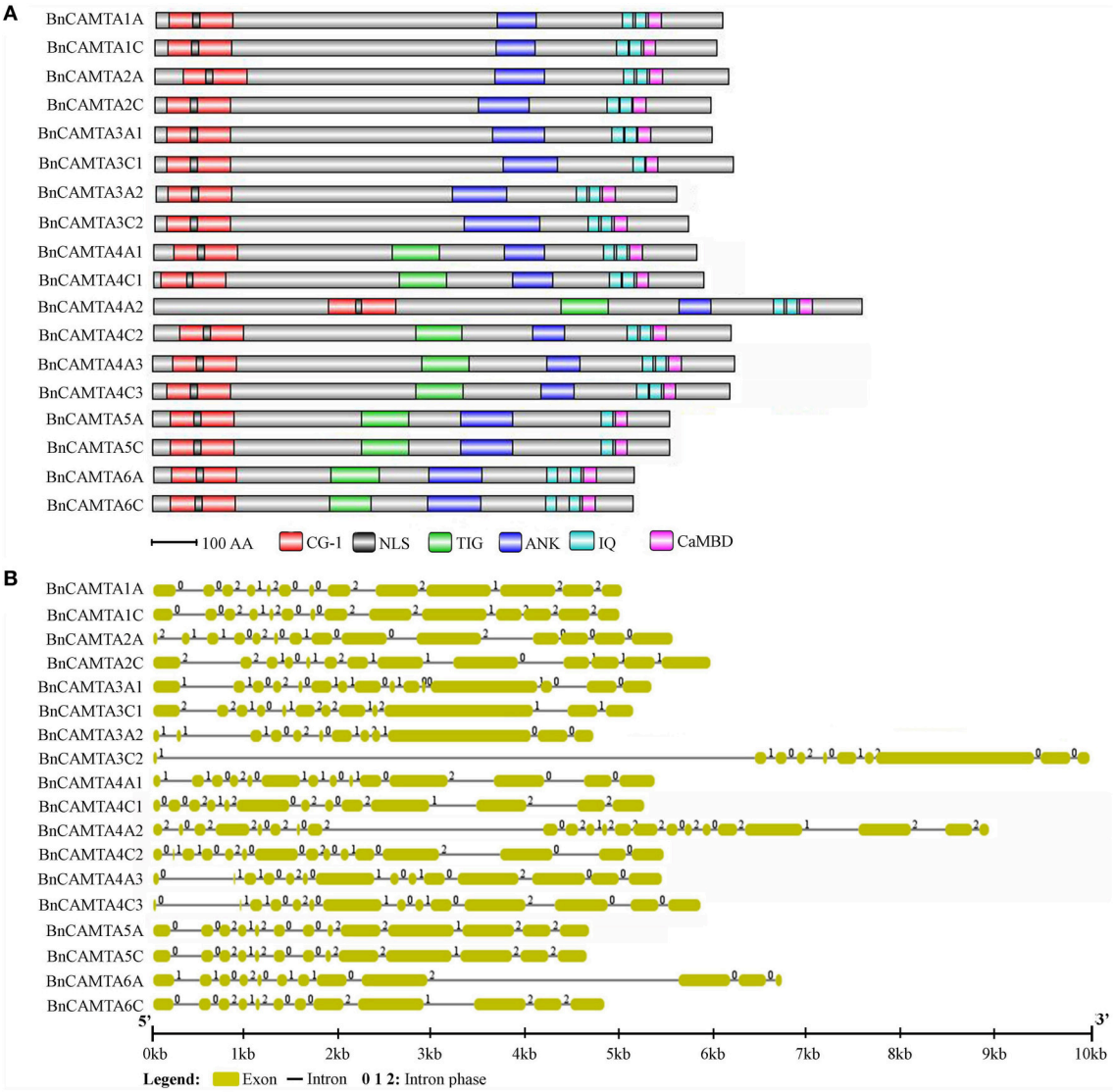


FIGURE 1 | Protein domain composition and gene structure of BnCAMTAs. (A) Schematic representation of domains of BnCAMTA proteins. Analyses of conserved domains were performed using the Pfam database (http://pfam.janelia.org/). NLS motifs were searched by Motif scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan). CaMBDs were analyzed using the Calmodulin Target Database (http://calcium.uhnres.utoronto.ca/ctdb/ctdb/). The domain structures of BnCAMTAs were drawn to scale using Domain Graph software (http://dog.biocuckoo.org/). Abbreviations: CG-1, CG-1 DNA-binding domain; NLS, nuclear localization signal motif; TIG, transcription-associated immuno globulin-like domain; ANK, ankyrin repeat domain; IQ, isoleucine glutamine motif; CaMBD, calmodulin-binding domain. **(B)** Exon-intron structure of *BnCAMTA* genes. The exons and introns are indicated by blue boxes gray lines, respectively. The *BnCAMTA* gene structures were drawn to scale using the Structure Display Server (GSDS, http://gsds.cbi.pku.edu.cn/).

proteins were predicted to contain a nuclear localization signal (NLS) in the N-terminus of the protein, consistent with their role as transcription factors that function in the nucleus (**Figure 1A**). This result indicated that the domain composition of CAMTAs in *B. napus* is similar to those in other plant species (Rahman et al., 2016).

Further, the exon-intron structure of the *BnCAMTA* genes was analyzed. The result demonstrated that the exon-intron configuration of most *BnCAMTA* genes was highly conserved with 11-14 introns, as observed for that of *CAMTA* genes in other plant species (Rahman et al., 2016). The exceptions were *BnCAMTA3C2* and *BnCAMTA4A2* genes. Both contained an intron with an unusual large size. Additionally, *BnCAMTA3C2* had only nine introns while *BnCAMTA4A2* possessed 21 introns (**Figure 1B**). Whether they exhibit distinct function from the others remains further study.

Conservation of CaMB Domain of BnCAMTAs

CaMB domain is indispensable to CAMTAs. To understand the conservation of this domain in BnCAMTAs, the corresponding sequence regions were aligned and compared with that in

well-studied Arabidopsis and tomato CAMTAs. The alignment revealed a conserved motif for functional residues as W X V X(2) L X K X(2) [LI] R W R X K X(3) [LF] [RKIV] X (**Figure 2**). Except for minor variation in some positions such as the 11^{th} and 21^{st} positions, the motif for BnCAMTAs generally fitted the one reported for Arabidopsis and tomato CAMTAs (W X V X(2) L X K X(2) [LF] R W R X [KR] X(3) [FL] R X). In this motif for BnCAMTAs, the 11^{th} hydrophobic residue was dominated by L except two sequences (BnCAMTAs 6A and 6C) as I. Similarly, the 21^{st} position was dominated by R, but it was K in BnCAMTAs 4A1, 4C1, 4A3, and 4C3 and I and V in BnCAMTAs 6A and 6C, respectively (**Figure 2**). Collectively, these data demonstrated that the motif of CaMB domain was highly conserved in CAMTA proteins of oilseed rape and other plant species.

Phylogenetic Relationship of *BnCAMTA* Genes

To gain insight into the phylogenetic relationship of *BnCAMTA* genes, a phylogenetic tree based on maximum-likelihood (ML) methods was constructed for 18 *B. napus* CAMTAs along with those from *B. rapa*, *B. oleracea*, Arabidopsis and tomato (**Table 1**, Table S2, and **Figure 3**). The phylogenetic analysis indicated that 18 BnCAMTAs clustered into three groups (I–III) with CAMTAs from other plant species with strong bootstrap support. All four memebers of BnCAMTA subgroups 5 and 6 constituted group I, together with CAMTA subgroups 5 and 6 from other plant species. All six members of BnCAMTA subgroup 4 gathered into group II along with CAMTA4s from other plant species, while all 8 non-TIG BnCAMTAs (all members of BnCAMTA subgroups 1, 2, and 3) formed group III, together with CAMTA subgroups 1, 2, and 3 from the other species (**Figure 3**). This phylogenetic tree revealed that all copies of CAMTAs in the two progenitors *B. rapa* and *B. oleracea* were well-inherited in *B. napus*. The similar clustering pattern was also obtained when the phylogenetic tree was reconstructed for the CAMTA proteins only from Arabidopsis and oilseed rape (Figure S2). It is noteworthy that different members of BnCAMTA subgroups 3 and 4 exhibited distinguishable phylogenetic distance to Arabidopsis CAMTA3 and CAMTA4. BnCAMTAs 3A1 and 3C1 were phylogenetically closer to AtCAMTA3 than BnCAMTAs 3A2 and 3C2. Similarly, BnCAMTAs 4A3 and 4C3 were phylogenetically closer to AtCAMTA4 than the other four BnCAMTA4s (Figure S2). These results indicated that *CAMTA3* and *CAMTA4* genes had been expanded in the three *Brassica* species compared with Arabidopsis although they belong to the same family (*Brassicaceae*). It is intriguing to probe whether members of the same subgroups function similarly or differentially considering that the pivotal role of AtCAMTA3 in plant defense has been unveiled.

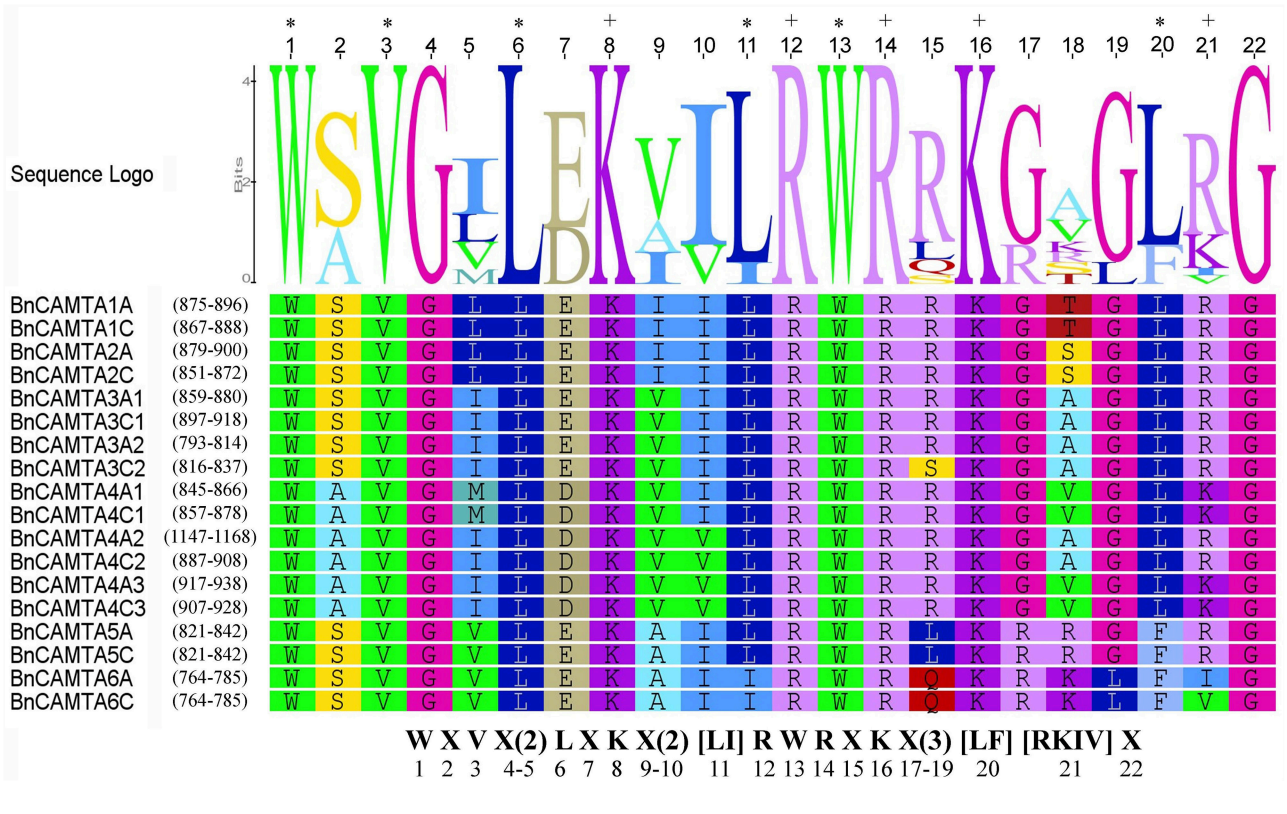


FIGURE 2 | Sequence logo of the CaMB domain of 18 BnCAMTA proteins. The "*" and "+" indicate conserved hydrophobic and positively charged residues in the CaMB domain. The conserved motif for these functional amino acids is shown at the bottom. The sequence logo of BnCAMTAs was drawn using Geneous R6 software (v6.1.6).

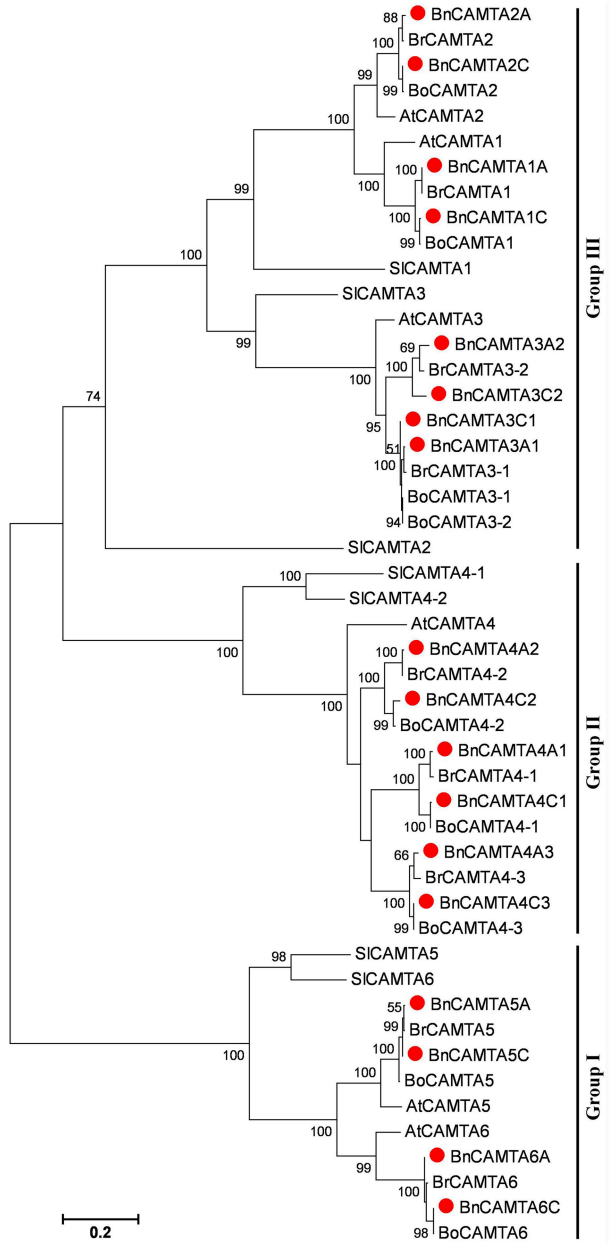


FIGURE 3 | Phylogenetic tree of oilseed rape CAMTAs along with homologs from its progenitors *B. rapa* and *B. oleracea* as well as Arabidopsis and tomato. Bootstrap values are displayed on the branches. Oilseed rape CAMTAs are marked with a solid red circle before the protein names. The tree was generated using the MEGA5 program by maximum likelihood (ML) methods.

Prediction of *cis*-Acting Elements in Promoters of *BnCAMTA* Genes

Nine well-defined and stress response-related *cis*-acting elements (DRE/CRT, ABRE, AuxRE, SARE, G- box, W-box, CG-box, P1BS, and SURE) were scanned in 1.5 kb sequences upstream of the ATG of *BnCAMTA* genes to obtain preliminary clues on how the *BnCAMTA* genes respond expressionally to stress stimuli. The results showed that there were various stress/stimulus response-related *cis*-acting elements in the promoter of *BnCAMTA* genes (**Table 2**). Analyses in both PLACE and PlantCARE databases predicted that *BnCAMTA* genes widely contained ABA responsive element (ABRE) and G-box element, some carried W-box element while a few possessed auxin responsive element (AuxRE) in their promoters (**Table 2**). Moreover, search in PLACE database predicted that some *BnCAMTA* genes owned additional *cis*-emelmets such as dehydration and cold responsive element (DRE/CRT), auxin responsive element (AuxRE), SA-responsive element (SARE), phosphate starvation-responsive element (P1BS), and sulfur-responsive element (SURE) in their promoters (**Table 2**). In addition, four *BnCAMTA* genes possessed 1–3 copies of CAMTA-recognizable CG-box elements according to the prediction result in the PLACE database, suggesting that CAMTAs might regulate their own gene transcription. Interestingly, every *BnCAMTA* gene contained at least one type of stress response-related *cis*-element, but the type of *cis*-element(s) in *BnCAMTA* genes was distinguishable (**Table 2**). Collectively, the stress-responsive *cis*-element analyses indicated that the *BnCAMTAs* are likely to be involved in plant response to various stresses and hormone signals.

Constitutive Expression of *BnCAMTA* Genes in Various Tissues of *B. napus*

To obtain a clue for the possible functions of the *BnCAMTA* genes, their expression profiles in different tissues or organs, including cotyledons of 1-week-old seedlings as well as roots, stems, and leaves of 4-week-old plants, were analyzed by semiquantitative RT-PCR. The results showed that different *BnCAMTA* genes exhibited distinct expression patterns. Seven out of 18 *BnCAMTA* genes (*BnCAMTAs* 2A, 4C1, 4C3, 5A, 5C, 6A, and 6C) were expressed highly in all investigated organs. Five genes (*BnCAMTAs* 1A, 2C, 3A1, 3C1, and 4A1) were expressed highly in stem, cotyledon and leaves but only weakly or even not in root. *BnCAMTA4C2* gene was expressed highly in stem but weakly in all other organs, while the remaining five *BnCAMTA* genes (*BnCAMTAs* 1C, 3A2, 3C2, 4A2, and 4A3) were only very weakly expressed in all types of organs (**Figure 4**), since their transcripts were detected only in the second round of RT-PCR using products of the first round PCR as template (Figure S3). Collectively, these expression data provided evidence to support that *BnCAMTA* genes play distinct roles in plant development.

Expression of *BnCAMTA* Genes in Response to Hormone and Chemical Treatments

To obtain a clue on functions of *BnCAMTA* genes, expressional response of these genes to multifunctional hormones SA and JA as well as oxalic acid (OA), the toxin secreted by the pathogen *S. sclerotiorum* was detected by RT-qPCR. From treatment perspective, SA strongly induced expression of *BnCAMTAs* 3A2, 3C2, and 4C1 by over 4 folds in at least one time point, and moderately induced expression of *BnCAMTAs* 4A1, all 5s and 6s by around 2 folds at the early time point (4 hpi) as well as

TABLE 2 | Predicted stress response-related *cis*-elements in the 1.5 kb sequence upstream of ATG of the *BnCAMTA* genes.

Gene Name	DRE/CRT		ABRE		AuxRE		SARE		G-box		W-box		CG-box		P1BS		SURE	
	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE	PLACE	PlantCARE
BnCAMTA1A	1						1				1				4		2	
BnCAMTA1C	4		1	1(6)	2		1		1	1(8)	2	2(6)	3		2			
BnCAMTA2A			1	4(6)					2	2(8)								
BnCAMTA2C	2		1	3(6)			3		3	3(6)	1		1					
BnCAMTA3A1				1(6)					1	1(6)					2			
BnCAMTA3C1				1(6)					1	1(6)	1	1(6)						
BnCAMTA3A2						1(7)					1				2		5	
BnCAMTA3C2						1(7)					2	1(6)	1				5	
BnCAMTA4A1									1	1(6)					2		1	
BnCAMTA4C1	2		1	1(6)	1				1	1(8)	1	1(6)					4	
BnCAMTA4A2	1		1	1(6)	1		2		1	1(6)	1						1	
BnCAMTA4C2			2	2(6)	1		1		2	2(6)								
BnCAMTA4A3			1	1(6)			1		1	1(6)	1		3					
BnCAMTA4C3			2	2(6)	3	2(6)	2				2						2	
BnCAMTA5A	1		2	4(6)	1		1		4	3(8)	1	2(6)					1	
BnCAMTA5C																	6	
BnCAMTA6A			1	1(7)							1	1(6)					1	
BnCAMTA6C	1		1	1(6)	1	1(5)			1	1(6)	1	1(6)			2			

The Matrix score (10 at maximum) for prediction in PlantCARE database is indicated in brackets. Abbreviations for cis-elements: DRE/CRT, dehydration and cold responsive element; ABRE, ABA-responsive element; AuxRE, auxin responsive element; SARE, SA-responsive promoter element; G-box, environmental signal response element; W-box, WRKY binding site; P1BS, phosphate starvation-responsive element; SURE, sulfur-responsive element; CG-box, CAMTA binding site.



FIGURE 4 | Constitutive expression patterns of *BnCAMTA* genes in various tissues. Expression patterns of the *BnCAMTA* genes in root (R), stem (S), cotyledon (C) and true leaf (L) were analyzed by semiquantitative RT-PCR. The oilseed rape β -Tubulin gene served as a loading control gene. The profile of electrophoresis on a 1.5% agarose gel of the products obtained from 32 cycles (28 cycles for control) of PCR was shown.

that of *BnCAMTA*s 1A and 4C2 at the late time point (24 hpi), but repressed expression of the remaining *BnCAMTA* genes. JA strongly induced expression of *BnCAMTA*s 3C2, 4A2 and all two 5s by over 4 folds at 12 hpi, and moderately induced expression of *BnCAMTA*s 1A, 1C, 2C, 3A1, 3A2, 4C2, and 6s at 12 and/or 24 hpi, but suppressed expression of the remaining *BnCAMTA* genes. OA generally induced expression of *BnCAMTA* genes at 24 hpi however, repressed expression of three subgroup 4 *BnCAMTA* genes (4A2, 4A3, and 4C3; **Figure 5**). From gene perspective, expression of most of the *BnCAMTA* genes was up-regulated by these three stimuli, although the level of alteration varied in response to different stimulus. However, expression of *BnCAMTA* genes 4A3 and 4C3 was significantly down-regulated by all stimuli. In addition, expression of *BnCAMTA* genes 1C, 2A, 2C, 3A1, 3C1, and 4A2 was reduced by SA, while expression of *BnCAMTA2A* and *BnCAMTA4A2* was repressed by JA and OA, respectively (**Figure 5**). The results indicated that the *BnCAMTA* genes widely but differentially respond at expression level to the three defense and stress-related signaling molecules SA, JA and OA.

Expression of *BnCAMTA* Genes during the Early Phase of *S. sclerotiorum* Infection

To probe the potential roles of BnCAMTAs in resistance to *S. sclerotiorum*, expression of the 18 *BnCAMTA* genes in oilseed rape leaves after *S. sclerotiorum* inoculation was inspected by qRT-PCR. The result showed that expression of six *BnCAMTA* genes (1A, 1C, 2A, 3A1, 3C1, and 6A) was significantly increased by over 2 folds after pathogen inoculation, peaking at 6 hpi except *BnCAMTA6A*, which reached a maximum at 12 hpi (**Figure 5**). Among them, *BnCAMTA3A1* and *BnCAMTA3C1* exhibited the most drastic change in expression. Their transcripts were increased by 9.1 and 7.0 folds, respectively, at 6 hpi. In addition, expression of six other *BnCAMTA* genes (2C, 3C2, 4A1, 4C2, 5A, and 6C) was also up-regulated in at least one time points but only at a change fold of less than 2.0. On the contrary, expression of four subgroup 4 *BnCAMTA* genes (4A2, 4A3, 4C1, and 4C3) and *BnCAMTA5C* was strongly decreased at the early time point of pathogen inoculation (6 hpi; **Figure 5**). These results confirmed that *BnCAMTA* genes are differentially transcriptionally responsive to *S. sclerotiorum* infection at the early phase.

Exogenous Supply of SA, JA, and OA Altered Resistance against *S. sclerotiorum* in Oilseed Rape Plants

The observation that some *BnCAMTA* genes are strongly responsive to SA, JA, and OA treatments and *S. sclerotiorum* inoculation prompted us to investigate the effect of these chemicals on resistance to *S. sclerotiorum* in oilseed rape plants. Leaves of oilseed rape plants were treated with these chemicals and inoculated with *S. sclerotiorum* at 4 h after treatments. As shown in **Figure 6**, SA and JA treatments obviously enhanced oilseed rape resistance to *S. sclerotiorum*, but OA treatment reduced plant resistance, as manifested by that *S. sclerotiorum* caused necrotic lesions were significantly smaller (1.4 and 1.6 cm at diameter) in SA- and JA-treated leaves, but larger (2.6 cm at diameter) in OA-treated leaves, than those (2.2 cm at diameter) in mock-inoculated control leaves at 36 hpi (**Figure 6**). This result indicated that SA and JA are associated with oilseed rape resistance to *S. sclerotiorum*.

Arabidopsis *CAMTA3* Negatively Regulated Resistance to *S. sclerotiorum*

To further explore the role of *CAMTA*s in plant resistance, we preformed inoculation analyses in six *Atcamta* mutants to examine their response to the devastating necrotrophic fungal pathogen *S. sclerotiorum*. Results of the inoculation analyses

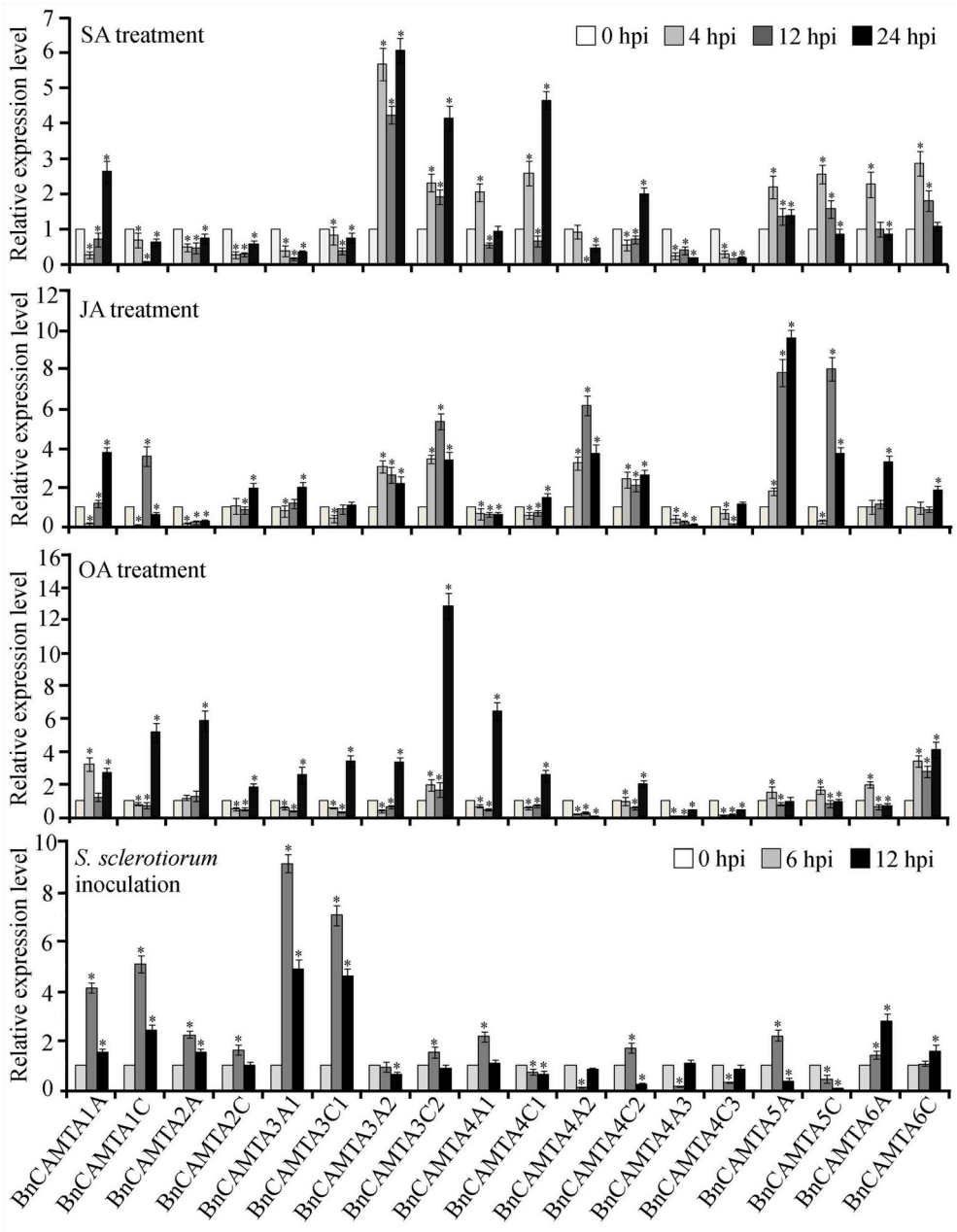


FIGURE 5 | Expression profiles of *BnCAMTA* genes in response to SA, JA and OA treatments and *S. sclerotiorum* inoculation. SA (salicylic acid, 1 mM), JA (jasmonic acid, 200 μM), and OA (oxalic acod, 1 mM)-sprayed and *S. sclerotiorum*-inoculated leaves of 4-week-old oilseed rape plants were sampled at the indicated time points for qRT-PCR analyses with β -Tubulin gene serving as a loading control gene. Data represent the mean ± SE of three independent experiments. Significant difference between mean values is indicated as a "*" ($p < 0.05$, DMRT).

showed that the *S. sclerotiorum* caused necrotic lesions in the *camta3* plants were significantly smaller (0.86 cm at diameter) than those in wild-type and the other mutant plants (over 1.17 cm at diameter) at 24 hpi (**Figure 7**), demonstrating that the *camta3* mutant plants were more resistant to *S. sclerotiorum* in comparison with wild-type and the other *camta* mutant plants. This result revealed that *CAMTA3* plays a negative role in plant resistance to *S. sclerotiorum*.

Arabidopsis *CAMTA3* Negatively Regulated Chitin-Elicited Accumulation of Hydrogen Peroxide

To provide some insights into the mechanisms of *AtCAMTA3* to regulate plant resistance, we inspected effect of AtCAMTA3 on accumulation of hydrogen peroxide induced by the PAMP chitin, which exists in the cell wall of the fungal pathogen *S. sclerotiorum*.

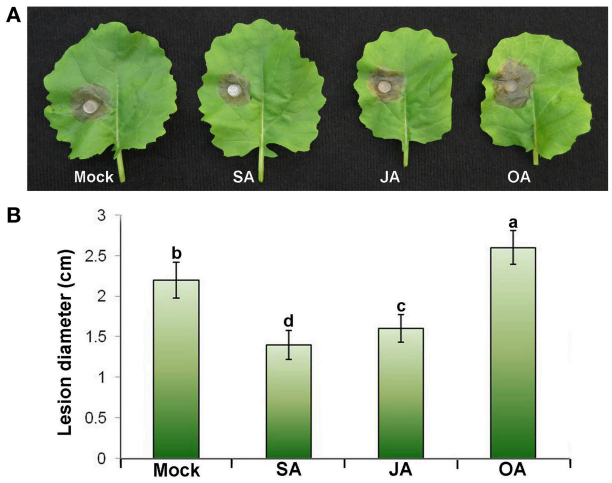


FIGURE 6 | Effect of exogenous treatment with SA, JA, and OA on resistance against *S. sclerotiorum* in oilseed rape plants. (A) The necrotic disease symptoms caused by *S. sclerotiorum* inoculation in leaves pretreated with 0.01% ethanol (mock), SA (1 mM), JA (200 μM), and OA (1 mM), respectively, at 4 h prior to pathogen inoculation. The photographs were taken at 36 hpi. **(B)** Statistical analysis of the lesion diameter. Data represent the mean ± SE of three independent experiments. Significant difference between mean values is indicated as small letters ($p < 0.05$, DMRT).

In response to 100 μg mL^{-1} chitin, *atcamta3* mutant plants accumulated much higher level of hydrogen peroxide, peaking at over 1200 RLU, than wild-type plants, culminating at about 600 RLU under current measuring system (**Figure 8**). This result demonstrated that *AtCAMTA3* negatively regulates chitin-triggered PTI as manifested by its negative regulation on chitin-triggered accumulation of hydrogen peroxide.

Arabidopsis *CAMTA3* Negatively Regulated the Expression of a Set of CGCG-Box Containing Defense Signaling Genes

To further elucidate the mechanisms of *AtCAMTA3* in regulating resistance to *S. sclerotiorum*, we examined the expression of four putative or confirmed *AtCAMTA3* targeted genes (*EDS1*, *NDR1*, *BAK1,* and *JIN1*) and three defense signaling pathway marker genes (*PR1*, *PDF1.2,* and *VSP1*) in wild-type and *Atcamta3* mutant plants before and after inoculating with *S. sclerotiorum*. These four genes were selected for this study because they are known to play important roles in plant resistance and PTI to *S. sclerotiorum* and/or other pathogens (Guo and Stotz, 2007; Du et al., 2009; Perchepied et al., 2010; Nie et al., 2012; Zhang et al., 2013; Macho and Zipfel, 2014). Meanwhile, *EDS1* and *NDR1* are targets of AtCAMTA3 (Du et al., 2009; Nie et al., 2012). Here, we found that the two PTI and/or *S. sclerotiorum*

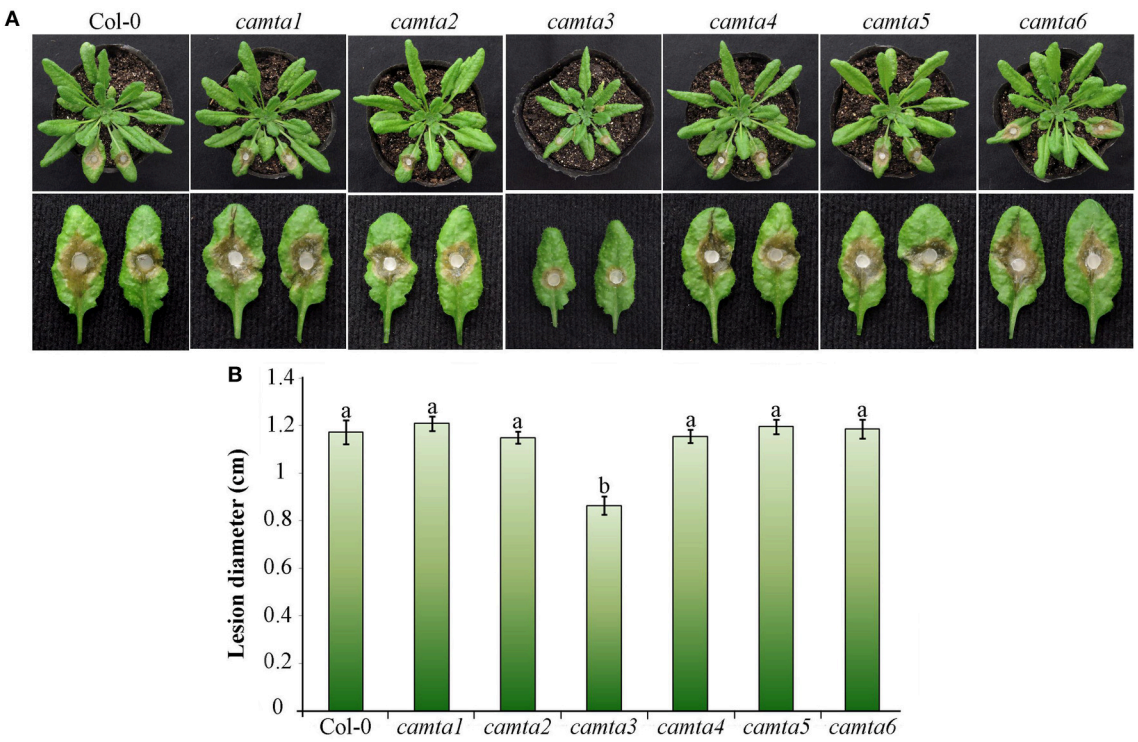


FIGURE 7 | *Atcamta3* mutant plants exhibited enhanced resistant to *S. sclerotiorum*. (A) The necrotic disease symptoms of the *Atcamta* mutants and Col-0 wild-type plants (upper panel) and detached leaves (lower panel) after inoculated with *S. sclerotiorum*. Photographs were taken at 30 hpi. **(B)** Statistical analysis of the lesion diameter. Data represent the mean ± SE of three independent experiments. Significant difference between mean values is indicated as small letters ($p < 0.05$, DMRT).

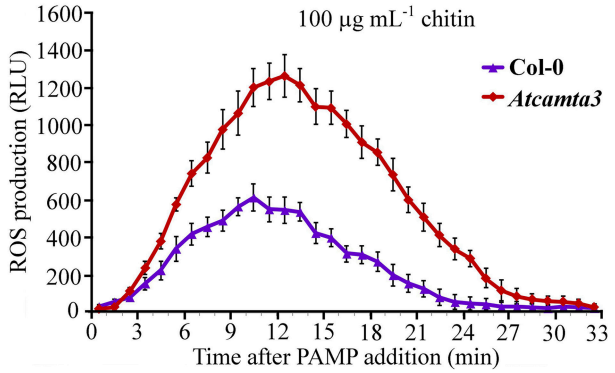


FIGURE 8 | *Atcamta3* mutant plants accumulated higher level of chitin-elicited H_2O_2. Dynamics of H_2O_2 accumulation in response to chitin elicitation in leaves of Col-0 and *Atcamta3* mutant plants was shown. Chitin-triggered H_2O_2 bursts were measured using luminol-based assay in leaf discs of Col-0 and *Atcamta3* mutant plants. Data are shown as relative luminal units (RLU) and represent the mean ± SE of three independent experiments.

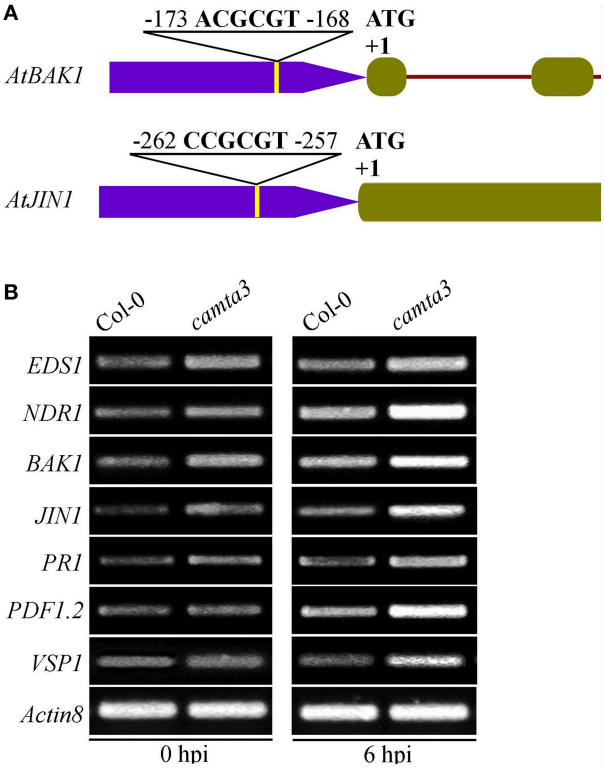


FIGURE 9 | Expression patterns of AtCAMTA3 target genes *EDS1* and *NDR1* and putative target genes *BAK1* and *JIN1* in *Atcamta3* mutant plants. (A) Schematic representation of CGCG elements in promoter regions of the *BAK1* and *JIN1* genes. **(B)** Expression patterns of four confirmed or putative AtCAMTA3-targeted genes and three defense signaling pathway marker genes (*PR1*, *PDF1.2*, and *VSP1*) in Col-0 wild-type and *Atcamta3* mutant plants at 0 and 6 h post *S. sclerotiorum* inoculation. Gene expression was examined by semiquantitative RT-PCR with the Arabidopsis ACTIN8 gene serving as a loading control gene. The profile of electrophoresis on a 1.5% agarose gel of the products obtained from 32 cycles (28 cycles for control) of PCR was shown.

resistance regulatory genes *BAK1* and *JIN1* also contained a CGCG *cis*-element in the region of −173 to −168 (ACGCGT) and −262 to −257 (CCGCGT), respectively, of their promoters (**Figure 9A**), they are therefore the potential targets of CAMTA3. Semiquantitative RT-PCR analysis revealed that the expression of *EDS1*, *NDR1*, *BAK1,* and *JIN1* in *Atcamta3* plants was obviously increased compared with the wild-type plants (**Figure 9B**), demonstrating that *AtCAMTA3* negatively regulates chitin-triggered immunity and resistance to *S. sclerotiorum* probably via negatively and directly regulating the expression of *EDS1*, *NDR1*, *BAK1,* and *JIN1*. Moreover, expression of *PR1*, *PDF1.2,* and *VSP1*, marker genes of SA, ethylene and JA defense signaling pathways was obviously higher in *Atcamta3* plants than in wild type plants (**Figure 9B**), indicating that *AtCAMTA3* negatively regulates resistance to *S. sclerotiorum* probably through modulating SA, ethylene and JA defense signaling pathways. In addition, in *Atcamta3* plants, transcripts of *EDS1*, *NDR1*, *BAK1,* and *JIN1* were still increased in response to *S. sclerotiorum* inoculation at 6 hpi (**Figure 9B**), suggesting that factor(s) other than AtCAMTA3 might respond to *S. sclerotiorum* inoculation to promote the expression of these defense signaling genes in *Atcamta3* plants.

DISCUSSION

Composition and Functions of *CAMTA* Gene Family in Oilseed Rape

In this study, we found that oilseed rape possesses a total of 18 CAMTAs. Oilseed rape is thus the plant species containing the highest number of CAMTAs among over 40 plant species whose *CAMTA* family has been identified to date (Bouché et al., 2002; Choi et al., 2005; Koo et al., 2009; Yang et al., 2012; Shangguan et al., 2014; Wang et al., 2015; Yang et al., 2015; Yue et al., 2015; Rahman et al., 2016). The number of *CAMTA* genes in oilseed rape is 3 folds as many as that in Arabidopsis, which is consistent with the ratio of total number of transcription factors in oilseed rape to that in Arabidopsis (Chalhoub et al., 2014). An important reason that oilseed rape carries so high number of CAMTAs is that oilseed rape is a tetraploid of the two progenitors *B. rapa* and *B. oleracea*, and thus contains copies of genes from both progenitors. As a matter of fact, we found that *B. rapa* and *B. oleracea* each own 9 CAMTAs, while the oilseed rape subgenomes A and C, which correspond to genomes of *B. rapa* and *B. oleracea*, respectively, each contain 9 CAMTAs with identical composition of subgroups as observed as *B. rapa* and *B. oleracea* (**Table 1**, Table S1, and **Figure 3**). Another reason that oilseed rape carries much higher number of CAMTAs than Arabidopsis is that oilseed rape genome has undergone *CAMTA* gene expansion compared with Arabidopsis genome. The subgenomes A and C of oilseed rape and genomes of *B. rapa* and *B. oleracea*, each contain 9 CAMTAs with 2 CAMTA3s and 3 CAMTA4s (**Table 1**, Table S1, and **Figure 3**), demonstrating that genomes of all three *Brassica* species have expanded *CAMTA3* and *CAMTA4* genes and this expansion in oilseed rape occurred before formation of oilseed rape. Expansion of *CAMTA3* and *CAMTA4* genes in *Brassica* species is unique

in all 11 species belonging to *Brassicaceae*, *Caricaece*, *Malvaceae*, *Rutaceae,* and *Myrtaceae* of Rosids whose *CAMTA* family has been identified to date (Figures 1, 2 in Rahman et al., 2016). The reason and significance of this expansion are unclear. While the function of CAMTA4 remains unknown, CAMTA3 in Arabidopsis has been well-recognized to play important roles in host and nonhost resistance against various pathogens (Galon et al., 2008; Du et al., 2009; Nie et al., 2012; Rahman et al., 2016). Therefore, whether both members of CAMTA3 in *B. rapa* and *B. oleracea* and all four in oilseed rape function in disease resistance is worthy of experimental clarification. In view of our comprehensive expression analyses, different members of subgroups BnCAMTA3 and BnCAMTA4 exhibit distinct expression profiles both constitutively in various tissues and in response to hormone treatments and pathogen inoculation. *BnCAMTA3A1* and *BnCAMTA3C1* are highly expressed in stem, cotyledon and true leaf while *BnCAMTA3A2* and *BnCAMTA3C2* are nearly not expressed in all these tissues (**Figure 4**). Meanwhile, *BnCAMTA3A1* and *BnCAMTA3C1* are not obviously responsive to SA and JA treatments but strongly responsive to *S. sclerotiorum* inoculation, while conversely, *BnCAMTA3A2* and *BnCAMTA3C2* are highly responsive to SA and JA treatments but not significantly responsive to *S. sclerotiorum* inoculation (**Figure 5**). Moreover, the gene structure of *BnCAMTA3C2* is distinct to the other members of *BnCAMTA3* subgroup (**Table 1**; **Figure 1B**). Similarly, distict expression patterns both constitutively in various tissues and in response to hormone treatments and pathogen inoculation are also observed for different members of *BnCAMTA4* genes (**Figures 4**, **5**). The gene structure of *BnCAMTA4A2* and the pI value of *BnCAMTA4A2* and *BnCAMTA4C2* are distinguishable from the other members of *BnCAMTA4* subgroup (**Table 1**; **Figure 1B**). Therefore, different members of subgroups BnCAMTA3 and BnCAMTA4 are most likely to play different roles in development, abiotic stress tolerance, and disease resistance. This seems to be also the case for functions of different subgroups of the *CAMTA* gene family in oilseed rape considering their distinct expression profiles both constitutively in various tissues and in response to diverse abiotic and biotic stimuli.

Role and Mechanism of *AtCAMTA3* in PTI and Resistance to the Necrotrophic Pathogen *S. sclerotiorum*

Role of CAMTAs in disease resistance against a wide range of biotrophic pathogens in various plants has been reported. These pathosystems include Arabidopsis against bacterial pathogens *Pst* DC3000 (Du et al., 2009) and *Xoo* (Rahman et al., 2016) as well as the fungal pathogen *Golovinomyces cichoracearum* (Nie et al., 2012), and rice against the bacterial pathogen *Xoo* and the fungal pathogen *Magnaporthe grisea* (Koo et al., 2009). However, Function of CAMTAs in plant disease resistance against necrotrophic pathogens has only been reported for one pathogen *Botrytis cinerea* (Galon et al., 2008; Li et al., 2014). In this study, using *camta* mutants, we demonstrate that AtCAMTA3 negatively regulates the resistance to the typical necrotrophic pathogen *S. sclerotiorum*, which is one of the most devastating fungal pathogens and causes the most imprtant disease, the white mold disease, in one of the most important oil-producing crops oilseed rape (Bolton et al., 2006). Additionally, oilseed rape *CAMTA* genes 1A, 1C, 3A1, and 3C1 are strongly responsive to *S. sclerotiorum* inoculation but differentially respond to the treatment with SA and JA, which play important roles in resistance to *S. sclerotiorum* (Guo and Stotz, 2007; Perchepied et al., 2010). Therefore, these four *CAMTA* genes may also play a role in resistance to *S. sclerotiorum* in oilseed rape. Taken together, these studies reveal that *CAMTA* genes, especially *CAMTA3*, contribute greatly to resistance against both biotrophic and necrotrophic pathogens in various plant species.

In addition, in this study, we provide some new intriguing points for the mechanisms of CAMTA3 to regulate PTI and *S. sclerotiorum* resistance. First, BAK1 might be the target of CAMTA3. BAK1 is a pivotal receptor kinase in PTI triggered by diverse PAMPs such as bacterial PAMP fig22 and fungal PAMP chitin (Macho and Zipfel, 2014). More importantly, it is also required for PTI triggered by SCFE1, a putative PAMP purified from *S. sclerotiorum* (Zhang et al., 2013). Interestingly, we found in this study that the *AtBAK1* gene contains a CGCG *cis*-element in the region of –173 to –168 (ACGCGT) of its promoter (**Figure 9A**). Furthermore, expression of *AtBAK1* is greatly enhanced in *Atcamta3* mutant plants compared with wild-type plants (**Figure 9B**). Moreover, *Atcamta3* mutant plants accumulate much higher level of chitin-elicited hydrogen peroxide than wild-type plants (**Figure 8**). Collectively, our results indicate that AtCAMTA3 negatively regulates the resistance to *S. sclerotiorum* probably via suppressing *AtBAK1*-meadited PTI. Second, CAMTA3 may target JIN1/MYC2 to directly modulate JA signaling thereby regulating plant defense against pathogens including *S. sclerotiorum*. JA signaling pathway is one of the most important plant defense pathways. This pathway is essential to the reisitance to *S. sclerotiorum* (Guo and Stotz, 2007; Perchepied et al., 2010). As a key component of JA signaling pathway, JIN1 is indispensible for the resistance to *S. sclerotiorum* (Guo and Stotz, 2007). We found that the *JIN1* gene contains a CGCG *cis*-element in the region of −262 to −257 (CCGCGT) of its promoter (**Figure 9A**). Further, expression of *AtJIN1* is greatly enhanced in *Atcamta3* mutant plants than in wild-type plants (**Figure 9B**). Together, our results suggest that AtCAMTA3 may modulate the JA signaling pathway via direct targeting JIN1 and thereby regulates the resistance to pathogens including *S. sclerotiorum*. In these scenarios, it will be very intriguing to confirm whether CAMTA3 can indeed direct bind and regulate expression of *BAK1* and *JIN1* by other approaches such as ChIP and EMSA assays. Finally, we oberved that expression of CAMTA3-targeted *EDS1* and *NDR1* genes is obviously increased in *Atcamta3* mutant plants than in wild-type plants (**Figure 9B**) as reported previously (Du et al., 2009; Nie et al., 2012; Rahman et al., 2016). These genes act upstream of SA signaling, which play a role in resistance to *S. sclerotiorum* (Guo and Stotz, 2007). Thus, *EDS1* and *NDR1* genes may also contribute to this resistance. The confirmation of function of these genes in this resistance will clarify the significance of CAMTA3-targeting of these two genes in resistance to *S. sclerotiorum*.

Based on our findings and the previously published reports (Benn et al., 2014; Rahman et al., 2016), we propose a schematic model for CAMTA3-mediated signaling in plants in response to pathogens and PAMPs (**Figure 10**). In this model, stimuli including pathogens such as *S. sclerotiorum* and *Xoo* as well as PAMPs such as chitin and flg22, may activate nucleotidyl cyclase (NC) to generate cyclic nucleotides including cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP), which activate Ca^{2+} channels such as cyclic nucleotide gated channels (CNGCs), leading to cytosolic Ca^{2+} influx (Qi et al., 2010; Ma and Berkowitz, 2011; Saand et al., 2015a,b). The cytosolic Ca^{2+} elevations are transduced by various Ca^{2+} sensor proteins including CaM, which activates CAMTA3. The activated CAMTA3 directly binds to the CGCG *cis*-elements in the promoter of defense-related target genes including *EDS1*, *NDR1*, *CBP60g*, *EIN3,* and *JIN1* and regulate their expression, which modulates the accumulation and signaling of SA, ET, and JA, and thereby alters disease resistance. Simultaneously, increased cytosolic Ca^{2+} would activate calcium-dependent protein kinases (CDPKs), which subsequently phosphorylate and activate RBOHD/F, resulting in ROS accumulation and thereby affecting hypersensitive response (HR) and plant disease resistance. Intriguingly, CAMTA3 may target *BAK1* to modulate the recognition complex, a beginning step for plant response to pathogens and PAMPs (**Figure 10**).

CONCLUSION

In the present study, we have identified and characterized 18 *CAMTA* genes in oilseed rape genome. They were inherited from the nine copies each in its progenitors *B. rapa* and *B. oleracea* and represented the highest number of CAMTAs in a given plant species identified to date. The oilseed rape CAMTAs clustered into three major groups and had expanded subgroups CAMTA3 and CAMTA4 uniquely in rosids species, which occurred before formation of oilseed rape. Comprehensive expression analyses indicated that *BnCAMTA* genes are likely to play distinct roles in development, abiotic stress tolerance and disease resistance. Among the four *BnCAMTA3* genes, *BnCAMTA3A1*

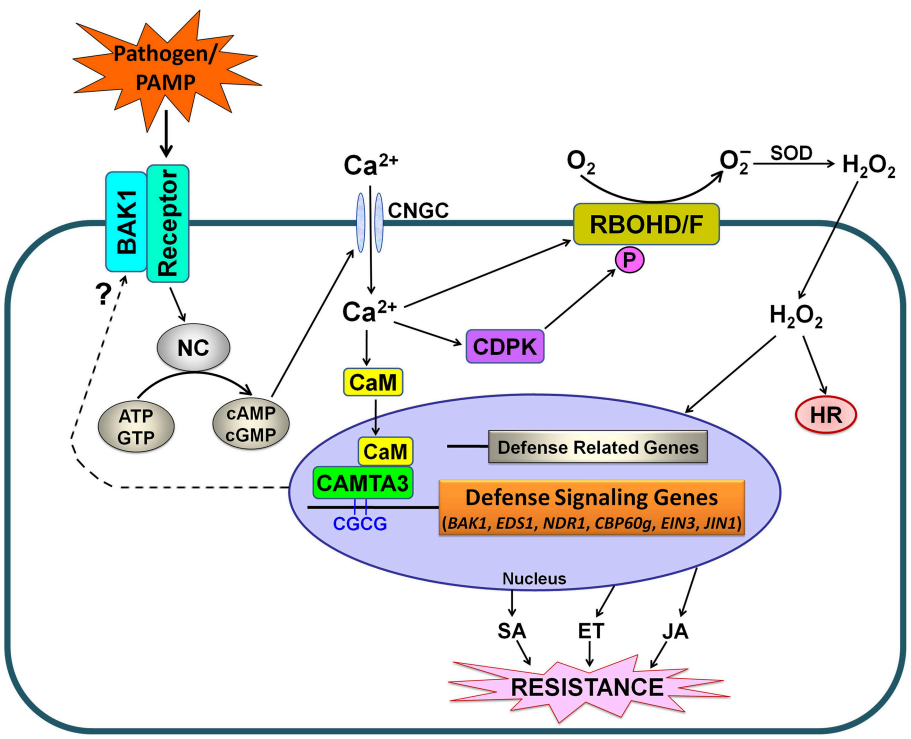


FIGURE 10 | Simplified schematic work model for CAMTA3-mediated defense signaling in plants. PAMPs from pathogens or exogenously supply are recognized by plant receptor proteins. Activation of the recognition complex generally requires BAK1. This recognition may lead to the activation of NC and consequently a rise in the concentration of cyclic nucleotides (cAMP and cGMP), resulting in cytosolic Ca^{2+} influx through opening of Ca^{2+} channels such as CNGCs. The cytosolic Ca^{2+} transients modulate CDPK-promoted and RBOHD/F-mediated ROS accumulation as well as CAMTA3-mediated denfense signaling. The activated CAMTA3 directly binds to the CGCG *cis*-elements in the promoter of defense-related target genes including *EDS1*, *NDR1*, *CBP60g*, *EIN3,* and *JIN1* and regulate their expression, which modulates the accumulation and signaling of SA, ET and JA, and thereby alters disease resistance. Additionally, CAMTA3 may target *BAK1* to mudulate the recognition complex, reflecting its global control of plant defense through regulating expression of the target genes at multiple nodes of the defense network. Abbreviations: BAK1, BRI1-associated receptor kinase; CaM, calmodulin; *CAMTA*, calmodulin binding transcriptional activator; CDPK, calcium-dependent protein kinases; CNGC, cyclic nucleotide gated channel; ATP, adenosine triphosphate; GTP, guanosine triphosphate; cAMP, cyclic adenosine monophosphate; cGMP, cyclic guanosine monophosphate; NC, nucleotidyl cyclase; HR, hypersensitive response; SOD, superoxide dismutase; SA, salicylic acid; ET, ethylene; JA, jasmonic acid.

and *BnCAMTA3C1* are most probably the functional homologs of *AtCAMTA3* and contribute to plant defense. Furthermore, functional analyses employing Arabidopsis *camta* mutants revealed that *CAMTA3* negatively regulates PAMP triggered immunity (PTI) probably by directly targeting *BAK1* and it also negatively regulates plant defense against pathogens such as *S. sclerotiorum* through suppressing JA signaling pathway probably via directly targeting *JIN1*. Our findings provide some insights into the composition of CAMTAs and their roles and functional mechanisms in plant defense.

AUTHOR CONTRIBUTIONS

HR and XZ conducted the bioinformatics and phylogenetic analyses. HR and YX carried out the gene expression and functional analysis, designed and analyzed all statistical data. XC conceived of the study, and participated in its design and coordination. XC and HR prepared the manuscript.

ACKNOWLEDGMENTS

We are grateful to Dr. Liquan Du, College of Life and Environmental Sciences, Hangzhou Normal University, China, for providing seeds of six Arabidopsis *CAMTA* knockout lines (*Atcamta1-6*). This work was financially supported by grants from the Genetically Modified Organisms Breeding Major Projects (no. 2014ZX0800905B), the National Natural Science Foundation of China (no. 31371892), the Zhejiang Provincial Natural Science Foundation of China (no. LZ12C14002) and the Special Fund for Agro-scientific Research in the Public Interest (no. 201103016).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00581

REFERENCES

Benn, G., Wang, C. Q., Hicks, D. K., Stein, J., Guthrie, C., and Dehesh, K. (2014). A key general stress response motif is regulated non-uniformly by *CAMTA* transcription factors. *Plant J.* 80, 82–92. doi: 10.1111/tpj.12620

Bolton, M. D., Thomma, B. P., and Nelson, B. D. (2006). *Sclerotinia sclerotiorum* (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Mol. Plant Pathol.* 7, 1–16. doi: 10.1111/j.1364-3703.2005.00316.x

Bouché, N., Scharlat, A., Snedden, W., Bouchez, D., and Fromm, H. (2002). A novel family of calmodulin-binding transcription activators in multicellular organisms. *J. Biol. Chem.* 277, 21851–21861. doi: 10.1074/jbc.M200268200

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

Choi, M. S., Kim, M. C., Yoo, J. H., Moon, B. C., Koo, S. C., Park, B. O., et al. (2005). Isolation of a calmodulin-binding transcription factor from rice (*Oryza sativa* L.). *J. Biol. Chem.* 280, 40820–40831. doi: 10.1074/jbc.M504616200

Doherty, C. J., Van Buskirk, H. A., Myers, S. J., and Thomashow, M. F. (2009). Roles for Arabidopsis *CAMTA* transcription factors in cold-regulated gene expression and freezing tolerance. *Plant Cell* 21, 972–984. doi: 10.1105/tpc.108.063958

Du, L., Ali, G. S., Simons, K. A., Hou, J., Yang, T., Reddy, A. S., et al. (2009). Ca^{2+}/calmodulin regulates salicylic-acid-mediated plant immunity. *Nature* 457, 1154–1158. doi: 10.1038/nature07612

Du, L., Yang, T., Puthanveettil, S., and Poovaiah, B. W. (2011). "Decoding of calcium signal through calmodulin: calmodulin-binding proteins in plants," in *Coding and Decoding of Calcium Signals in Plants*, ed S. *Luan (Berlin; Heidelberg: Springer)*, 177–233.

Finkler, A., Ashery-Padan, R., and Fromm, H. (2007). CAMTAs: calmodulin-binding transcription activators from plants to human. *FEBS Lett.* 581, 3893–3898. doi: 10.1016/j.febslet.2007.07.051

Galon, Y., Finkler, A., and Fromm, H. (2010). Calcium-regulated transcription in plants. *Mol. Plant* 3, 653–669. doi: 10.1093/mp/ssq019

Galon, Y., Nave, R., Boyce, J. M., Nachmias, D., Knight, M. R., and Fromm, H. (2008). Calmodulin-binding transcription activator *(CAMTA)* 3 mediates biotic defense responses in Arabidopsis. *FEBS Lett.* 582, 943–948. doi: 10.1016/j.febslet.2008.02.037

Guo, A. Y., Zhu, Q. H., Chen, X., and Luo, J. C. (2007). GSDS: a gene structure display server. *Yichuan* 29, 1023–1026. doi: 10.1360/yc-007-1023

Guo, X., and Stotz, H. U. (2007). Defense against *Sclerotinia sclerotiorum* in Arabidopsis is dependent on jasmonic acid, salicylic acid, and ethylene signaling. *Mol. Plant Microbe Interact.* 20, 1384–1395. doi: 10.1094/mpmi-20-11-1384

Kim, Y., Park, S., Gilmour, S. J., and Thomashow, M. F. (2013). Roles of *CAMTA* transcription factors and salicylic acid in configuring the low-temperature transcriptome and freezing tolerance of Arabidopsis. *Plant J.* 75, 364–376. doi: 10.1111/tpj.12205

Koo, S. C., Choi, M. S., Chun, H. J., Shin, D. B., Park, B. S., Kim, Y. H., et al. (2009). The calmodulin-binding transcription factor OsCBT suppresses defense responses to pathogens in rice. *Mol. Cells* 27, 563–570. doi: 10.1007/s10059-009-0081-4

Kudla, J., Batistic, O., and Hashimoto, K. (2010). Calcium signals: the lead currency of plant information processing. *Plant Cell* 22, 541–563. doi: 10.1105/tpc.109.072686

Laluk, K., Prasad, K. V., Savchenko, T., Celesnik, H., Dehesh, K., Levy, M., et al. (2012). The calmodulin-binding transcription factor SIGNAL RESPONSIVE1 is a novel regulator of glucosinolate metabolism and herbivory tolerance in Arabidopsis. *Plant Cell Physiol.* 53, 2008–2015. doi: 10.1093/pcp/pcs143

Li, X., Huang, L., Zhang, Y., Ouyang, Z., Hong, Y., Zhang, H., et al. (2014). Tomato SR/CAMTA transcription factors SlSR1 and SlSR3L negatively regulate disease resistance response and SlSR1L positively modulates drought stress tolerance. *BMC Plant Biol.* 14:286. doi: 10.1186/s12870-014-0286-3

Ma, W., and Berkowitz, G. A. (2011). Ca^{2+} conduction by plant cyclic nucleotide gated channels and associated signaling components in pathogen defense signal transduction cascades. *New Phytol.* 190, 566–572. doi: 10.1111/j.1469-8137.2010.03577.x

Macho, A. P., and Zipfel, C. (2014). Plant PRRs and the activation of innate immune signaling. *Mol. Cell* 54, 263–272. doi: 10.1016/j.molcel.2014.03.028

Nie, H., Zhao, C., Wu, G., Wu, Y., Chen, Y., and Tang, D. (2012). SR1, a calmodulin-binding transcription factor, modulates plant defense and ethylene-induced senescence by directly regulating NDR1 and EIN3. *Plant Physiol.* 158, 1847–1859. doi: 10.1104/pp.111.192310

Perchepied, L., Balagué, C., Riou, C., Claudel-Renard, C., Rivière, N., Grezes-Besset, B., et al. (2010). Nitric oxide participates in the complex interplay of defense-related signaling pathways controlling disease resistance to *Sclerotinia sclerotiorum* in *Arabidopsis thaliana*. *Mol. Plant Microbe Interact.* 23, 846–860. doi: 10.1094/mpmi-23-7-0846

Qi, Z., Verma, R., Gehring, C., Yamaguchi, Y., Zhao, Y., Ryan, C. A., et al. (2010). Ca^{2+} signaling by plant *Arabidopsis thaliana* Pep peptides depends on AtPepR1, a receptor with guanylyl cyclase activity, and cGMP-activated Ca^{2+} channels. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21193–21198. doi: 10.1073/pnas.1000191107

Qiu, Y., Xi, J., Du, L., Suttle, J. C., and Poovaiah, B. W. (2012). Coupling calcium/calmodulin-mediated signaling and herbivore-induced plant response through calmodulin-binding transcription factor AtSR1/CAMTA3. *Plant Mol. Biol.* 79, 89–99. doi: 10.1007/s11103-012-9896-z

Rahman, H., Yang, J., Xu, Y. P., Munyampundu, J. P., and Cai, X. Z. (2016). Phylogeny of plant CAMTAs and role of AtCAMTAs in nonhost resistance to *Xanthomonas oryzae* pv. *oryzae*. *Front. Plant Sci.* 7:177. doi: 10.3389/fpls.2016.00177

Reddy, A. S., Ali, G. S., Celesnik, H., and Day, I. S. (2011). Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. *Plant Cell* 23, 2010–2032. doi: 10.1105/tpc.111.084988

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., and Yao, X. (2009). DOG 1.0: illustrator of protein domain structures. *Cell Res.* 19, 271–273. doi: 10.1038/cr.2009.6

Saand, M. A., Xu, Y. P., Li, W., Wang, J. P., and Cai, X. Z. (2015a). Cyclic nucleotide gated channel gene family in tomato: genome-wide identification and functional analyses in disease resistance. *Front. Plant Sci.* 6:303. doi: 10.3389/fpls.2015.00303

Saand, M. A., Xu, Y. P., Munyampundu, J. P., Li, W., Zhang, X. R., and Cai, X. Z. (2015b). Phylogeny and evolution of plant cyclic nucleotide-gated ion channel (CNGC) gene family and functional analyses of tomato CNGCs. *DNA Res.* 22, 471–483. doi: 10.1093/dnares/dsv029

Shangguan, L., Wang, X., Leng, X., Liu, D., Ren, G., Tao, R., et al. (2014). Identification and bioinformatic analysis of signal responsive/calmodulin-binding transcription activators gene models in *Vitis vinifera*. *Mol. Biol. Rep.* 41, 2937–2949. doi: 10.1007/s11033-014-3150-5

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Wang, G., Zeng, H., Hu, X., Zhu, Y., Chen, Y., Shen, C., et al. (2015). Identification and expression analyses of calmodulin-binding transcription activator genes in soybean. *Plant Soil* 386, 205–221. doi: 10.1007/s11104-014-2267-6

Yang, T., Peng, H., Whitaker, B. D., and Conway, W. S. (2012). Characterization of a calcium/calmodulin-regulated SR/CAMTA gene family during tomato fruit development and ripening. *BMC Plant Biol.* 12:19. doi: 10.1186/1471-2229-12-19

Yang, T., Peng, H., Whitaker, B. D., and Jurick, W. M. (2013). Differential expression of calcium/calmodulin-regulated SlSRs in response to abiotic and biotic stresses in tomato fruit. *Physiol. Plant.* 148, 445–455. doi: 10.1111/ppl.12027

Yang, T., and Poovaiah, B. W. (2002). A calmodulin-binding/CGCG box DNA-binding protein family involved in multiple signaling pathways in plants. *J. Biol. Chem.* 277, 45049–45058. doi: 10.1074/jbc.M207941200

Yang, Y., Sun, T., Xu, L., Pi, E., Wang, S., Wang, H., et al. (2015). Genome-wide identification of *CAMTA* gene family members in *Medicago truncatula* and their expression during root nodule symbiosis and hormone treatments. *Fronti. Plant Sci.* 6:459. doi: 10.3389/fpls.2015.00459

Yue, R., Lu, C., Sun, T., Peng, T., Han, X., Qi, J., et al. (2015). Identification and expression profiling analysis of calmodulin-binding transcription activator genes in maize (*Zea mays* L.) under abiotic and biotic stresses. *Fronti. Plant Sci.* 6:576. doi: 10.3389/fpls.2015.00576

Zhang, L., Du, L., Shen, C., Yang, Y., and Poovaiah, B. W. (2014). Regulation of plant immunity through ubiquitin-mediated modulation of Ca^{2+}-calmodulin-AtSR1/CAMTA3 signaling. *Plant J.* 78, 269–281. doi: 10.1111/tpj.12473

Zhang, W., Fraiture, M., Kolb, D., Löffelhardt, B., Desaki, Y., Boutrot, F. F. G., et al. (2013). Arabidopsis RECEPTOR-LIKE PROTEIN30 and receptor-like kinase SUPPRESSOR OF BIR1-1/EVERSHED mediate innate immunity to necrotrophic fungi. *Plant Cell* 25, 4227–4241. doi: 10.1105/tpc.113.117010

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Rahman, Xu, Zhang and Cai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 April 2016
doi: 10.3389/fpls.2016.00451

Growth and Metal Accumulation of an *Alyssum murale* Nickel Hyperaccumulator Ecotype Co-cropped with *Alyssum montanum* and Perennial Ryegrass in Serpentine Soil

Catherine L. Broadhurst[1,2]* *and Rufus L. Chaney*[3]

[1] *Environmental Microbiology and Food Safety Laboratory, U.S. Department of Agriculture Agricultural Research Service, Beltsville, MD USA,* [2] *Department of Food Science and Nutrition, University of Maryland, College Park, MD, USA,* [3] *Crop Systems and Global Change Laboratory, U.S. Department of Agriculture Agricultural Research Service, Beltsville, MD, USA*

The genus *Alyssum* (Brassicaceae) contains Ni hyperaccumulators (50), many of which can achieve 30 g kg^{-1} Ni in dry leaf. Some *Alyssum* hyperaccumulators are viable candidates for commercial Ni phytoremediation and phytomining technologies. It is not known whether these species secrete organic and/or amino acids into the rhizosphere to solubilize Ni, or can make use of such acids within the soil to facilitate uptake. It has been hypothesized that in fields with mixed plant species, mobilization of metals by phytosiderophores secreted by Graminaceae plants could affect *Alyssum* Ni, Fe, Cu, and Mn uptake. We co-cropped the Ni hyperaccumulator *Alyssum murale*, non-hyperaccumulator *A. montanum* and perennial ryegrass in a natural serpentine soil. All treatments had standard inorganic fertilization required for ryegrass growth and one treatment was compost amended. After 4 months *A. murale* leaves and stems contained 3600 mg kg^{-1} Ni which did not differ significantly with co-cropping. Overall Ni and Mn concentrations were significantly higher in *A. murale* than in *A. montanum* or *L. perenne*. Copper was not accumulated by either *Alyssum* species, but *L. perenne* accumulated up to 10 mg kg^{-1}. *A. montanum* could not compete with either *A. murale* or ryegrass, and neither *Alyssum* species survived in the compost-amended soil. Co-cropping with ryegrass reduced Fe and Mn concentrations in *A. murale* but not to the extent of either increasing Ni uptake or affecting plant nutrition. The hypothesized *Alyssum* Ni accumulation in response to phytosiderophores secreted by co-cropped grass did not occur. Our data do not support increased mobilization of Mn by a phytosiderophore mechanism either, but the converse: mobilization of Mn by the *Alyssum* hyperaccumulator species significantly increased Mn levels in *L. perenne*. Tilling soil to maximize root penetration, adequate inorganic fertilization and appropriate plant densities are more important for developing efficient phytoremediation and phytomining approaches.

Keywords: ***Alyssum murale, Lolium perenne,*** **nickel hyperaccumulators, ryegrass, co-cropping, phytoremediation, phytomining**

OPEN ACCESS

Edited by:
Sarvajeet Singh Gill, Maharshi Dayanand University, India

Reviewed by:
Mirza Hasanuzzaman, Sher-e-Bangla Agricultural University, Bangladesh
Narsingh Chauhan, Maharshi Dayanand University, India

***Correspondence:**
Catherine L. Broadhurst
leigh.broadhurst@ars.usda.gov

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *21 December 2015*
Accepted: *22 March 2016*
Published: *08 April 2016*

Citation:
Broadhurst CL and Chaney RL (2016) Growth and Metal Accumulation of an Alyssum murale Nickel Hyperaccumulator Ecotype Co-cropped with Alyssum montanum and Perennial Ryegrass in Serpentine Soil. Front. Plant Sci. 7:451. doi: 10.3389/fpls.2016.00451

INTRODUCTION

More than 400 plant species are known to naturally accumulate high levels of metals such as Cd, Cu, Co, Mn, Ni, and Zn (Baker et al., 2010; Krämer, 2010; van der Ent et al., 2013). The genus *Alyssum* (Brassicaceae) contains the greatest number of reported Ni hyperaccumulators (50), many of which can achieve 30 g kg^{-1} Ni in dry leaf biomass (Baker and Brooks, 1989; Reeves and Adigüzel, 2008; van der Ent et al., 2015). Previously we have demonstrated commercially feasible phytoremediation and phytomining technologies that can potentially clean up Ni-contaminated soils and recover high purity Ni metal (Chaney et al., 1999, 2010; Li et al., 2003a,b; Nkrumah et al., 2016). The technology employs the Ni-hyperaccumulating species *Alyssum murale* and *A. corsicum* to phytoextract Ni from a range of Ni-rich soil types. *A. murale* and *A. corsicum* are endemic to serpentine soils developed from ultramafic rock throughout Mediterranean Southern Europe.

Ni localization patterns have been determined for 10 *Alyssum* Ni hyperaccumulator species/ecotypes (Krämer et al., 1997; Psaras et al., 2000; Küpper et al., 2001; Kerkeb and Krämer, 2003; Marmiroli et al., 2004; Broadhurst et al., 2004a,b, 2009; McNear et al., 2005; Asemaneh et al., 2006; Tappero et al., 2007). Nickel is stored mainly in the leaves, and is particularly concentrated in in vacuoles of epidermal cells and trichome pedicels. *Alyssum* hyperaccumulators also accumulate appreciable Mn in the same locations that contain Ni (Broadhurst et al., 2004b, 2009).

Although Ni hyperaccumulation is a constitutive property for these *Alyssum* species, it is not known whether they secrete organic and/or amino acids into the rhizosphere to solubilize Ni, or can make use of such acids within the soil to greatly facilitate uptake. Other than rhizobiome interactions, there is essentially no evidence for unusual ligand species or highly elevated ligand concentrations associated with Ni in *Alyssum* (McNear et al., 2010; Centofanti et al., 2013). There is evidence that rhizosphere bacteria endemic to serpentine soils may stimulate Ni uptake and this may be an important factor explaining why field trials and native vegetation consistently outperform pot and hydroponic studies with respect to phytoextraction yields (Abou-Shanab et al., 2003, 2007; Rajkumar et al., 2009, 2013; Cabello-Conejo et al., 2014; Visioli et al., 2015). Two serpentine-endemic bacteria in particular (*Microbacterium arabinogalactanolyticum* and *M. oxydans*) were shown to strongly increase Ni accumulation in *A. murale* (Abou-Shanab et al., 2003, 2007). Similarly, endemic *Arthrobacter sp*. increased Ni uptake in *A. pintodasilvae* and *A. serpyllifolium* (Cabello-Conejo et al., 2014).

Cd/Zn hyperaccumulators have not shown evidence for specialized ligand secretion into the rhizosphere either (Zhao et al., 2001; Whiting et al., 2001a; Sterckeman et al., 2005; Wang et al., 2006). Specifically, root exudates collected from the Cd/Zn hyperaccumulator *Noccaea caerulescens* F.K. Mey (Brassicaceae) (syn. *Thlaspi caerulescens* J &C Presl) did not mobilize Cd, Cu, Fe, or Zn (Zhao et al., 2001). Further, Cd/Zn hyperaccumulators may not take advantage of potential phytosiderophore-related improvements in metal solubilization provided by intercropping with Graminaceae.

The grass family of plants differs from all other plant families by using a different mechanism of absorbing Fe from soils. All other species use a combination of acidification of the rhizosphere and reduction of ferric to ferrous coupled with absorption of ferrous ion. Instead, Graminaceae use a combination of chelating amino acids, the phytosiderophores of the mugineic acid family of compounds, for specific uptake of intact Fe–phytosiderophore chelates. It is known that phytosiderophores are not highly specific to Fe and can increase mobilization and possibly support uptake of Zn, Mn, Ni, Cu, and Cd as well (Zhang et al., 1991a,b; Marschner and Römheld, 1994; Awad and Römheld, 2000). Intercropping peanut (*Arachis hypogaea* L.), for example, with maize, oats, barley or wheat significantly increased Fe, Cu, and Zn uptake to the extent that Fe deficiency in peanut could be mitigated (Zuo and Zhang, 2011).

Previous results from co-cropping hyperaccumulators and grasses are mixed. *N. caerulescens* had no increase in Cd or Zn concentration when grown in the same pot with ryegrass (*Lolium perenne* L.), but yield was almost doubled in an experiment where plants were grown with sufficient time and soil volume to establish potential rhizosphere interactions with or without root mingling (Jiang et al., 2010). It was determined that the ryegrass did not solubilize Cd and Zn, however Fe was not discussed. The improved yield could be at least partially due to improved Fe availability. Increased P availability from arbuscular mycorrhizal fungi which are known to colonize Graminaceae including ryegrass (Grimold et al., 2005) is another factor which could significantly affect yield since many of the metalliferous soils that hyperaccumulators are native to are P deficient. However, co-planting the Cd/Zn hyperaccumulator *Sedum alfredii* (Hance) with ryegrass reduced both *S. alfredii* yield and Cd uptake (Wang et al., 2013). Co-planting with corn improved *S. alfredii* yield by providing shade but did not increase Cd uptake. Cd and Zn uptake by corn was unaltered by co-cropping and corn did not suffer phytotoxicity (Wu et al., 2007).

Co-cropping barley (*Hordeum vulgare* L.) and *N. caerulescens* in multiple metal-rich soils from a biosolids management facility showed little evidence for interaction between plants other than a slight increase in Cd, Cu, Ni, Zn in co-cropped pots with root interaction vs. *N. caerulescens* alone, but this probably reflected simple depletion of metals in the relatively small soil volume utilized, and not a specific phytosiderophore mechanism (Gove et al., 2002). Again, Fe was not considered in the experiment. Both Gove et al. (2002) and Jiang et al. (2010) observed an increase in ryegrass Cd but not Zn concentrations when grown with *N. caerulescens*. Similarly, Whiting et al. (2001a) showed no interaction between *N. caerulescens* and *Festuca rubra* L. with respect to Zn levels or yield.

Improved growth and reduced Zn uptake by the non-hyperaccumulator *Thlaspi arvense* L. was reported when *T. arvense* and *T. caerulescens* were grown together in pots that allowed root intermingling. Zinc salts were added to the soils at a level that was phytotoxic to *T. arvense*. Zn hyperaccumulation by *T. caerulescens* was not affected, however yield increased when root intermingling was allowed, leading to the conclusion that this system could facilitate revegetation of contaminated soils (Whiting et al., 2001b).

Herein we report a co-cropping experiment with *Alyssum* hyperaccumulator and non-hyperaccumulator species and perennial ryegrass in a natural serpentine soil. The soil is infertile and high in Ni, but is not Ni phytotoxic (Zhang et al., 2007) and supports native vegetation. Soils such as this are candidates for Ni phytomining (Chaney et al., 2010; Nkrumah et al., 2016). We tested whether ryegrass facilitates Ni, Fe, and Mn uptake by *Alyssum,* whether co-cropping with ryegrass affects *Alyssum* yield, and whether *Alyssum* hyperaccumulator and non-hyperaccumulator species can benefit from co-cropping.

MATERIALS AND METHODS

Horticulture

A. murale (Waldst. et Kit) "Kotodesh," a Ni-hyperaccumulator, was grown from seed collected from a wild Albanian serpentine population. All *Alyssum* hyperaccumulator species known have leaves covered with stellate trichomes (**Figure 1**). Nickel is stored in leaf epidermal cells, particularly in the trichome pedicels. *A. montanum* L. "Mountain Gold" (a non-hyperaccumulator species that also has leaf trichomes) was grown from commercial seed (Hazzard's Seeds, Deford, MI). Both *Alyssum* species were started in flats with Promix® potting soil and standard fertilization (half-strength Miracle Grow®). After 40 days healthy *Alyssum* seedling roots were rinsed to remove potting medium and transplanted to prepared soils in pots. Four weeks after transplant, when seedlings had become established, commercial perennial ryegrass (*L. perenne* L. "Amazing GS," Ampac Seed, Tangent, OR) was seeded directly into the pots. Twenty cm polyethylene pots which hold about 3 kg air dry soil were utilized. Plastic mesh was not used across the pot over the drain holes in order to avoid interference with root growth. All watering was with deionized water, and plastic trays were placed under each pot. To avoid overwatering *Alyssum,* 250–650 ml was added 2 or 3 times per week to ensure that soil dried between waterings. The co-cropped plants were grown for an additional 12 weeks.

We utilized Brockman variant serpentine soil from Josephine Co., Oregon (Typic Xerochrepts), air dried and sieved <4 mm using stainless steel sieves. Standard inorganic fertilization for serpentine soils (75 mg N as NH_4NO_3, 100 mg P as KH_2PO_4, 500 mg Ca as $CaSO_4 \cdot 2H_2O$, and 0.5 mg B as H_3BO_3 per kg soil) was added. The Brockman soil as collected is pH 6.6, very high in Ni (4710 mg kg^{-1}), more than adequate in Mn and Fe, but deficient in Ca and P (**Table 1**). One set of treatments had 30 vol% (10%DW) aged dairy manure compost from the USDA Beltsville composting facility mixed into the serpentine soil. Fertilizer rates allowed normal growth of ryegrass on this soil which would not normally support growth of non-serpentinophytes.

The experiment was conducted in the USDA Beltsville greenhouse under controlled temperature and light conditions and ambient humidity. Photoperiod was 15/9 h day/night. During this time supplemental high-intensity sodium and incandescent lights capable of supplying 400 $\mu mol\, m^{-2} s^{-1}$ supplemented sunlight if necessary. Daytime temperature was 24°C with cooling initiated at 27°C. Nighttime temperature was 18°C with cooling initiated at 21°C. During the final 3 weeks of growth in late May and June supplemental lighting was turned off to avoid overheating.

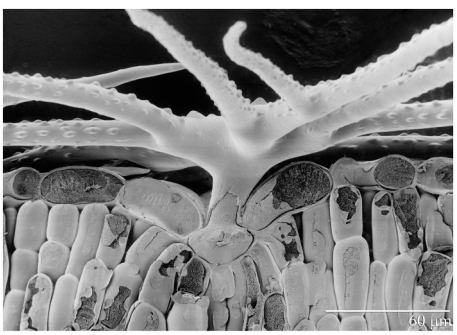


FIGURE 1 | Scanning electron photomicrograph of frozen hydrated Ni-hyperaccumulator *Alyssum murale* 'Kotodesh' and surface of leaf. Note stellate trichome; Ni is very high in the vacuole below the trichome.

Experimental Treatments

Two types of soil and six planting schemes made up 12 treatments, with three replicates per treatment. *Alyssum* plants that died soon after transplanting were replaced for the first 2 weeks of growth. Overall *Alyssum* grew 4 months after transplant, and rye grass grew 3 months after seeding. At harvest plant roots filled the pot and intermingled but plants were not pot-bound.

Soil A: Natural Brockman variant serpentine soil.
Soil B: Natural Brockman variant serpentine soil with 10 wt% compost.

TABLE 1 | Representative average Brockman soil parameters over 12 years testing.

Analyte	Units	Result
pH		6.30
organic matter	%	3.1
Clay	%	21.0
Sand	%	58.0
Fe	g kg^{-1}	222
Mn	g kg^{-1}	3.88
Ni	g kg^{-1}	4.71
Co	mg kg^{-1}	354
Cu	mg kg^{-1}	45.0
Zn	mg kg^{-1}	180
Ca, exchangeable	meq/100 g	1.0
Mg, exchangeable	meq/100 g	6.2
Ni, exchangeable	mg kg^{-1}	32.0
Ni, DPTA extractable	mg kg^{-1}	124
P, Bray extractable	mg kg^{-1}	0.02

Note the very low Ca and P levels and Ca:Mg ratio characteristic of serpentine soils.

Planting Scheme:

1. A. murale, four plants per pot spread radially and evenly.
2. A. montanum, four plants per pot spread radially and evenly.
3. *L. perenne* four groups of approximately 20 seedlings spread radially and evenly.
4. *A. murale* two plants, and *A. montanum*, 2 plants with the same plant spread as #1, alternating species.
5. *A. murale* two plants, and *L. perenne* two groups of approximately 20 seedlings, alternating species.
6. *A. montanum* two plants, and *L. perenne* two groups of approximately 20 seedlings alternating species.

Plant Material Metals Analysis

All clean, healthy aerial plant material was harvested. Material that was stained from the red serpentine soil or unhealthy was discarded. Harvested plant material was washed in a dilute detergent bath and rinsed in deionized water to remove adhering soil particles. Plants were dried for 72 h at 60°C, weighed, and ashed in a 480°C oven for 16 h. After cooling, the ash was digested with 2 ml concentrated HNO_3, mixed well and then heated to dryness. The sample was then dissolved in 10 ml 3 *N* HCl, filtered through Whatman #40 filter paper and brought to volume in a 25 ml volumetric flask using 0.1 *N* HCl. Concentrations of Ca, Cd, Cu, Fe, K, Mg, Ni, Mn, P, and Zn were determined by inductively-coupled plasma atomic emission spectrometry using 40 mg L^{-1} yttrium as an internal standard in all samples and standard solutions.

Soil Analysis

Total soil metals were measured by atomic absorption spectrometry after digestion with boiling HNO_3. Exchangeable Ca, Mg, and Ni were obtained by extracting 5 g air-dried soil with 50 mL 1.0 M ammonium acetate at pH 7, soil texture by pipette method, and organic matter by combustion. The Bray-2 method was used to estimate phytoavailable P. The DTPA-extraction used 5 g soil per 50 mL standard DTPA extractant rather than the usual 10 g per 20 mL because of the high metals levels in this and other Ni-rich soils studied in our laboratory.

RESULTS

Yields and dry weight metal concentrations are reported in **Table 2**, and examples of co-cropped healthy plants in the serpentine soil (**A**) are given in **Figures 2**, **3**. All results were statistically analyzed by ANOVA with SAS. None of the *Alyssum* transplants survived in compost-amended soil **B**. The compost evidently contained pathogen(s) that both species were susceptible to, and it also kept the soil damp longer between waterings. Normally this is desirable in pot studies, however *A. murale* in particular is adapted to semi-arid conditions, and once established survives with watering once per week or less. The symptoms exhibited were consistent with fungal infection. None of the plants in soil **A** or seedlings in Promix were affected.

In soil **A**, *A. murale* shoots contained approximately 3600 mg kg^{-1} Ni which did not differ significantly with co-cropping (**Figure 4**). Overall Ni and Mn concentrations were significantly higher in *A. murale* than *A. montanum* or *L. perenne* (**Figures 4**, **5**). However, *A. murale* Fe concentrations were significantly reduced ($p < 0.05$) by co-cropping with ryegrass, and Mn was somewhat reduced ($p < 0.4$) but half-pot yield was equivalent. In general Fe concentrations were unreliable in *A. montanum* due to contamination with Fe^{3+} oxide staining deep within the leaves, coupled with only a small amount of plant material available for analysis, but this does not affect the

TABLE 2 | Half-pot dry weight yield and element concentrations of all treatments that survived.

Species	Treatment	Yield g ½ pot^{-1}	Ca G kg^{-1}	Cu mg kg^{-1}	Fe mg kg^{-1}	K g kg^{-1}	Mg g kg^{-1}	Mn mg kg^{-1}	Ni mg kg^{-1}	P g kg^{-1}	Zn mg kg^{-1}
A. murale	A1 alone	7.5 ± 3.2	13.6 ± 0.6	2.35 ± 0.42	336 ± 260	14.4 ± 1.2	2.09 ± 0.28	225 ± 150	3320 ± 330	2.92 ± 0.37	27.9 ± 7.6
	A4 + mon	10.3 ± 5.1	13.1 ± 2.4	2.31 ± 0.27	285 ± 240	15.4 ± 1.1	2.08 ± 0.75	202 ± 19	4030 ± 1170	3.32 ± 0.57	20.7 ± 1.4
	A5 + peren	7.87 ± 0.75	12.4 ± 1.3	2.01 ± 0.51	89.3 ± 40	14.1 ± 1.9	2.41 ± 0.62	145 ± 32	3640 ± 830	2.50 ± 0.54	18.1 ± 6.4
A. montanum	A2 alone	3.60 ± 0.4	24.1 ± 2.4	1.76 ± 1.1	589 ± 350	19.9 ± 2.6	7.50 ± 1.1	47 ± 10	50 ± 16	3.54 ± 0.55	14.1 ± 2.1
	A4 + mur*	1.33	22.0	1.20	342	22.2	6.76	64	57.5	3.54	13.5
	A6 + peren	0.71 ± 0.30	26.0 ± 1.5	2.00 ± 0.37	1689 ± 80	25.9 ± 6.2	9.32 ± 1.3	54.7 ± 21	71.3 ± 14	3.87 ± 0.53	16.8 ± 2.7
L. perenne	A3 alone	2.7 ± 0.7	4.37 ± 0.3	6.46 ± 0.06	179 ± 20	25.7 ± 2.5	4.46 ± 0.51	110 ± 28	50.7 ± 4.9	3.95 ± 0.46	29.6 ± 1.5
	A5 + mur	1.75 ± 1.2	4.71 ± 0.84	5.93 ± 0.86	231 ± 120	26.3 ± 5.1	4.78 ± 0.67	202 ± 15	39.7 ± 28	2.96 ± 0.56	28.3 ± 3.2
	A6 + mon	4.01 ± 0.95	4.10 ± 0.62	5.49 ± 1.6	352 ± 350	20.5 ± 6.9	4.38 ± 0.77	102 ± 28	45.7 ± 2.3	2.79 ± 0.34	24.4 ± 3.9
	B3 alone	3.1 ± 1.7	2.45 ± 0.35	8.79 ± 0.92	60.0 ± 00	37.1 ± 5.6	3.49 ± 0.09	52.7 ± 8.1	19.3 ± 4.0	4.00 ± 0.74	82.5 ± 15
	B5 + mur	8.61 ± 2.8	2.25 ± 0.65	10.1 ± 0.97	62.0 ± 27	41.8 ± 11	3.24 ± 0.31	46 ± 8.5	23.3 ± 2.5	3.61 ± 0.70	77.1 ± 5.2
	B6 + mon	7.20 ± 0.79	2.81 ± 0.38	11.1 ± 0.89	96.7 ± 31	46.9 ± 7.4	3.43 ± 0.18	53.3 ± 6.0	19 ± 2.0	3.71 ± 0.32	80.5 ± 2.0

Mean and s.d. of three replicates except where noted. Soil A: 100% Brockman variant serpentine; soil B Brockman serpentine soil with 10 dw% manure compost mixed in. Plants designated "alone" grown in monoculture; other treatments were mixed culture as listed. Only L. perenne survived in the manure compost amended soil.
**only two replicates survived.*



FIGURE 2 | *A. murale* (most of the plant material in pot, with long stems, and darker oblong leaves) and *A. montanum (*a few plants in center with larger, lighter leaves in a rosette pattern) co-cropped in the fertilized serpentine soil. *A. montanum* grew less vigorously than *A. murale* in the fertilized serpentine soil, and was outcompeted and shaded after 4–5 weeks.



FIGURE 3 | *Lolium perenne and A. montanum* co-cropped in the fertilized serpentine soil. The nonhyperacumulator *Alyssum* species was not able to compete with the grass.

relationship between Ni and Mn in *A. murale* vs. *A. montanum*. Although Fe is much higher in *A. montanum* than *A. murale*, both Ni and Mn are significantly lower ($p < 0.01$), a result that cannot be due to soil contamination.

Nickel concentrations in *A. montanum* and *L. perenne* remained relatively constant and below 75 mg kg^{-1} in for **A2** through **A6**. However, there was increased variability of the Ni concentration in both *A. murale* or *L. perenne* with *A. murale* intercropping, and intercropping with *A. murale* significantly increased Mn in *L. perenne* (**Figure 5**). Calcium concentrations were three to six times greater in *Alyssum* species as compared to *L. perenne* due to the high Ca in *Alyssum* leaf trichomes (**Table 2**; **Figure 1**). However, Cu concentrations were consistently greater in *L. perenne* than *Alyssum* (**Figure 6**).

In soil **A**, *A. montanum* could not compete with either *A. murale* or ryegrass and was nearly killed by co-cropping, with 10-fold yield reductions (**Figures 3**, **4**). Due to its poor growth, *A. montanum* did not significantly affect the growth or metal concentrations of *A. murale* or *L. perenne*. However, there was a significant ryegrass yield reduction with *A. murale* co-cropping. Ryegrass yield was increased when co-cropped with *A. montanum* because it thoroughly out-competed *A. montanum* with only half the plants.

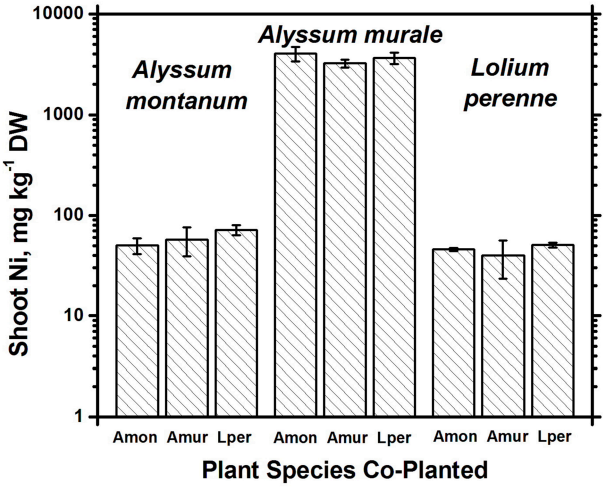


FIGURE 4 | Ni concentrations for species in monoculture and co-cropped. Nickel levels were significantly greater in the hyperacccumulator *Alyssum murale* ($p < 0.001$) but did not differ significantly within a given species as a function of co-cropping. Error bars means ± Standard Error.

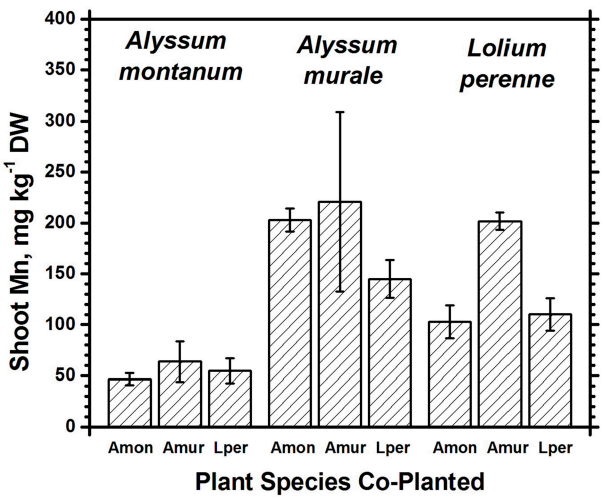


FIGURE 5 | Mn concentrations for species in monoculture and co-cropped. Mn levels were significantly greater in *A. murale* than *A. montanum*. Mn was increased in *L. perenne* with *A. murale* co-cropping. Error bars means ± Standard Error.

Because *Alyssum* did not survive in the treatments with compost, all three ryegrass planting schemes grew ryegrass only. Essentially there were nine replicates for **B3**, all of which had Ni concentrations that did not differ significantly from one another, but did differ from **A3** and **A5**, as expected due to the high levels of Ni in the serpentine soil (**Figure 4**). There was a trend for increased yield with compost but it was not significant. Calcium, Fe, Mg, and Ni concentrations in ryegrass were reduced with compost and Cu and Zn concentrations were increased. Ni concentrations in ryegrass were about twice as high in treatment **A** compared to **B**, while the reverse was true for Cu and Zn.

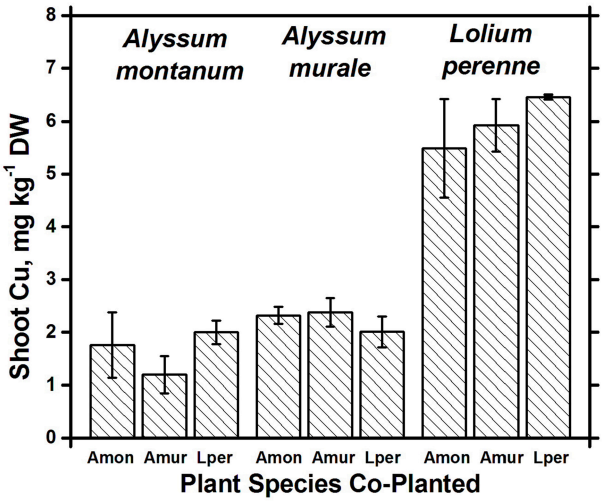


FIGURE 6 | Cu concentrations for species in monoculture and co-cropped. Copper levels were significantly greater in *Lolium perenne* compared to both *Alyssum* species ($p < 0.001$). With the exception of a slight Cu reduction in *A. montanum* co-cropped with *A. murale*, Cu concentrations did not differ significantly within a given species as a function of co-cropping. Error bars means ± Standard Error.

DISCUSSION

Our results indicate there is no value with respect to phytomining or phytoextraction in co-cropping *A. murale* with *L. perenne*. Neither yield nor Ni uptake was improved; the ryegrass shoots only interfered with the growth of *A. murale*. The average full-pot yield for *A. murale* grown alone was 15.0 ± 6.3 g, therefore phytoextraction could be doubled just with *Alyssum* monoculture. Further, given larger pot sizes or field growth, *A. murale* and *A. corsicum* develop extensive root systems than can increase shoot Ni concentration up to five times that achieved in 1 kg pots (Baklanov et al., 2015; Bani et al., 2015a). Therefore, root growth interference from the equally extensive *L. perenne* root system is almost certainly a negative factor with respect to maximizing Ni phytoextraction. Co-cropping *Lupinus albus* and *A. murale* in natural serpentine soils showed similar results in a study investigating whether co-cropping with a nitrogen-fixing plant could improve overall *A. murale* Ni phytoextraction (Jiang et al., 2015). Without supplemental P fertilization, 90% of the biomass in the pots was *L. albus*. With P fertilization *A. murale* increased to 39%, however Ni accumulation in the shoots was significantly reduced compared to monocropping. Overall Ni phytoextraction was maximized in the monocrop with P fertilization.

Co-cropping with ryegrass somewhat reduced Fe and Mn concentrations in *A. murale* but not to the extent of either increasing Ni uptake or affecting plant nutrition, so the result, while interesting, has a neutral effect on phytoextraction in this soil. The increased variability of the Ni concentration in both *A. murale* and *L. perenne*, and increased Mn in *L. perenne* with *A. murale* co-cropping may reflect increased rhizosphere mobilization of Ni and Mn by *A. murale* but in this experiment it did not translate to any tangible benefit. The hypothesized increase in Ni accumulation in response to phytosiderophores secreted by co-cropped grasses clearly did not occur. Our data do not support increased mobilization of Mn by a phytosiderophore mechanism either, but the converse: mobilization of Mn by the *Alyssum* hyperaccumulator species significantly increased Mn levels in the grass.

A. montanum could not compete with either *A. murale* or ryegrass and was nearly killed by co-cropping. In field growth it would be unlikely to survive. In contrast to results with *Noccaea* (Whiting et al., 2001b), there would be no value in utilizing an *Alyssum* hyperaccumulator to improve the growth of a non-hyperaccumulator; *A. montanum* yield was strongly reduced by co-cropping, yet it grew well alone in the fertilized serpentine soil.

A murale and *A. montanum* accumulated about 13 and 24 g Ca kg^{-1} respectively, consistent with all previous observations in which $CaCO_3$ nodules cover the surface of the trichomes (Krämer et al., 1997; Psaras et al., 2000; Küpper et al., 2001; Kerkeb and Krämer, 2003; Marmiroli et al., 2004; Broadhurst et al., 2004a,b; Broadhurst et al., 2009; McNear et al., 2005; Asemaneh et al., 2006; Tappero et al., 2007). Calcium fertilization was necessary for *L. perenne* growth in this experiment, however *Alyssum* hyperaccumulator species native to typically low Ca, low Ca:Mg ratio serpentine soils nonetheless accumulate Ca in the absence of fertilization. *A. montanum* is not a Ni hyperaccumulator but had twice the Ca of *A. murale,* but only half the Mn.

Both Ni and Mn concentrations were significantly higher in *A. murale* than *A. montanum* or *L. perenne*. The high variability in the *A. murale* Mn concentration is typical of the species, which in natural serpentine soils is observed to hyperaccumulate Mn only in some leaves on a given plant. If Mn soil levels are exceedingly high without addition of Ni, Mn is not hyperaccumulated throughout the plant and instead becomes phytotoxic (Broadhurst et al., 2004b, 2009; Tappero et al., 2007). Ni hyperaccumulators are very specific to Ni and to a lesser extent Mn and Co, and do not non-selectively accumulate/hyperaccumulate other transition metals such as Fe, Cr, or Cu. In the case of Cu, despite 3600 mg kg^{-1} Ni accumulation, *A. murale* Cu concentrations were only 2 mg kg^{-1}, far below *L. perenne*, which accumulated typical foliar Cu levels for ryegrass. These observations support a specific relationship between Mn accumulation and Ni hyperaccumulation (Broadhurst et al., 2009; Ghaderian et al., 2015) rather than a general situation for *Alyssum* species where Mn uptake and storage is related to enhanced Ca uptake to synthesize the unique trichome tissues (McNear and Kupper, 2013).

Although the compost utilized was a standard, mature aged product from USDA Beltsville it was very detrimental to *Alyssum* growth, most likely due to pathogenic fungi. We have repeatedly observed fungal infections in *Alyssum* species grown in humid summer greenhouse conditions. Although we grew the plants in the late winter/spring season and in a majority of their native soil, they were nonetheless unable to survive transplant to the manure compost amended soil. Seeding in the pot was tried but the germination rate of *A. montanum* was below 40% and seedlings that did come up were very weak. Several plants were

transplanted to soil **B** and grown outdoors but also succumbed with the same disease pattern. However, both *Alyssum* species grew very well alone in the fertilized serpentine soil with no evidence of disease or phytotoxicity.

In a similar recent study, Álvarez-López et al. (2016) grew the hyperaccumulators *A. serpyllifolium ssp. lusitanicum*, *A. serpyllifolium ssp. malacitanum*, *A. pintodasilvae,* and *A. bertolonii*. in their native serpentine soil with 2.5, 5, and 10 wt% commercial municipal solid waste compost added. These species grow slowly so did not achieve the large shrub size that *A. murale* can in one season. The lower levels of compost addition significantly increased yield but no further benefits were achieved with 10%. All levels of compost addition reduced extractable Ni; at the 10% level the reduction was 11-fold. Overall yield was lower without compost but with inorganic NPK fertilization, however NPK addition did not affect Ni accumulation. In an Albanian field trial with ultramafic Vertisols, Bani et al. (2015b) found *A. murale* yield was increased 10-fold with 120 kg NPK and 77 kg Ca ha^{-1} plus monocot herbicide to control Graminaceae—as opposed to encouraging co-cropping. These agronomic practices increased Ni phytoextraction yield from 2.0 to 29.5 kg ha^{-1}. Thus, in a long-term field Ni phytoextraction or phytomining situation, standard inorganic fertilization may be both adequate and preferable. If a compost source is utilized, it should be tested with every species/ecotype used in the field program prior to application. Another factor to consider is a possible negative effect of compost biota on serpentine-endemic rhizobacteria which can act to facilitate Ni uptake. The two bacteria shown to strongly increase Ni accumulation in *A. murale* (*M. arabinogalactanolyticum* and *M. oxydans*) were isolated from the Oregon soil that we utilized in this study (Abou-Shanab et al., 2003, 2007), thus were potentially present in each pot. They may not have thrived in the compost-amended soil just as the *Alyssum* species did not, however rhizobiome interactions cannot explain our results.

Ryegrass growth was not negatively affected by the compost and the Ni concentration was significantly reduced without inducing Fe or Mn deficiency. With fertilization and adequate water ryegrass grew reasonably well on the serpentine soil; adding compost would be a significant benefit to retain soil moisture and improve root growth. Ryegrass yield may have increased in treatment **B** if grass was cut one or two times during the experiment. This was not done because in a field intercropping situation *A. murale* would need to grow as long as the season permits in order to maximize Ni phytoextraction, and it would not be practicable to selectively cut the ryegrass. Similarly, regular light irrigation and cool, relatively humid conditions to maximize *L. perenne* yield would not be practicable since *A. murale* grows better with infrequent but thorough waterings and relatively hot, sunny, low humidity conditions. In commercial phytomining of Ni, weed control to prevent grasses would normally be practiced to limit competition for water and nutrients (Bani et al., 2015b).

Overall, tilling soil to maximize root penetration, adequate inorganic fertilization and appropriate plant densities are more important for developing efficient phytoremediation and phytomining approaches with *Alyssum* Ni hyperaccumulator species than organic soil amendments or co-cropping.

AUTHOR CONTRIBUTIONS

CB Principal Experimentalist, analyzed data, and wrote ms. RC Research Group Leader, involved in all experimentation in the laboratory. Co-designed experiment and co-wrote ms.

FUNDING

RC is a Federal Employee, US Department of Agriculture Agricultural Research Service.

CB is posted at US Department of Agriculture Agricultural Research Service and is a State of Maryland Employee.

REFERENCES

Abou-Shanab, R. A., Angle, J. S., Delorme, T. A., Chaney, R. L., van Berkum, P., Moawad, H., et al. (2003). Rhizobacterial effects on nickel extraction from soil and uptake by *Alyssum murale. New Phytol.* 158, 219–224. doi: 10.1046/j.1469-8137.2003.00721.x

Abou-Shanab, R. A., van Berkum, P., and Angle, J. S. (2007). Heavy metal resistance and genotypic analysis of metal resistance genes in gram-positive and gram-negative bacteria present in Ni-rich serpentine soil and in the rhizosphere of *Alyssum murale*. *Chemosphere* 68, 360–367. doi: 10.1016/j.chemosphere.2006.12.051

Álvarez-López, A., Prieto-Fernández, Á., Cabello-Conejo, M. I., and Kidd, P. S. (2016). Organic amendments for improving biomass production and metal yield of Ni-hyperaccumulating plants. *Sci. Total Environ.* 548, 370–379. doi: 10.1016/j.scitotenv.2015.12.147

Asemaneh, T., Ghaderian, S. M., Crawford, S. A., Marshall, A. T., and Baker, A. J. M. (2006). Cellular and subcellular compartmentation of Ni in the Eurasian serpentine plants *Alyssum bracteatum*, *Alyssum murale* (Brassicaceae) and *Cleome heratensis* (Capparaceae). *Planta* 225, 193–202. doi: 10.1007/s00425-006-0340-y

Awad, F., and Römheld, V. (2000). Mobilization of heavy metals from contaminated calcareous soils by plant-born, microbial and synthetic chelators and their uptake by wheat plants. *J. Plant Nutr.* 23, 1847–1855. doi: 10.1080/01904160009382147

Baker, A. J. M., and Brooks, R. R. (1989). Terrestrial higher plants which hyperaccumulate metallic elements: a review of their distribution, ecology and phytochemistry. *Biorecovery* 1, 81–126.

Baker, A. J. M., Ernst, W. H. O., Van Der Ent, A., Malaisse, F., and Ginocchio, R. (2010). "Metallophytes: the unique biological resource its ecology, and conservational status in Europe, central Africa and Latin America," in *Ecology of Industrial Pollution*, eds L. C. Batty and K. B. Hallberg (Cambridge, UK: University Press), 7–40. doi: 10.1017/cbo9780511805561.003

Baklanov, I., Ryan, T. C., Chaney, R. L., and Davis, A. P. (2015). "Effect of soil volume on yield of and Ni hyperaccumulation by *Alyssum corsicum,*" in *Abstracts International Phytotechnologies Society Conference* (Manhattan, KS).

Bani, A., Echevarria, G., Sulçe, S., and Morel, J. L. (2015b). Improving the agronomy of *Alyssum murale* for extensive phytomining: a five-year field study. *Int. J. Phytorem.* 17, 117–127. doi: 10.1080/15226514.2013.862204

Bani, A., Echevarria, G., Zhang, X., Benizri, E., Laubie, B., Morel, J. L., et al. (2015a). The effect of plant density in nickel-phytomining field experiments with *Alyssum murale* in Albania. *Aust. J. Bot.* 63, 72–77. doi: 10.1071/bt14285

Broadhurst, C. L., Chaney, R. L., Angle, J. A., Erbe, E. F., and Maugel, T. K. (2004a). Nickel localization and response to increasing Ni soil levels in leaves of the Ni hyperaccumulator *Alyssum murale*. *Plant Soil* 265, 225–242. doi: 10.1007/s11104-005-0974-8

Broadhurst, C. L., Chaney, R. L., Angle, J. A., Maugel, T. K., Erbe, E. F., and Murphy, C. A. (2004b). Simultaneous hyperaccumulation of nickel, manganese and calcium in *Alyssum* leaf trichomes. *Environ. Sci. Technol.* 38, 5797–5802. doi: 10.1021/es0493796

Broadhurst, C. L., Tappero, R. V., Maugel, T. K., Erbe, E. F., Sparks, D. L., and Chaney, R. L. (2009). Interaction of nickel and manganese in accumulation and localization in leaves of the Ni hyperaccumulators *Alyssum murale* and *Alyssum corsicum*. *Plant Soil* 314, 35–48. doi: 10.1007/s11104-008-9703-4

Cabello-Conejo, M. I., Becerra-Castro, C., Prieto-Fernández, A., Monterroso, C., Saavedra-Ferro, A., Mench, M., et al. (2014). Rhizobacterial inoculants can improve nickel phytoextraction by the hyperaccumulator *Alyssum pintodasilvae*. *Plant Soil* 379, 35–50. doi: 10.1007/s11104-014-2043-7

Centofanti, T., Sayers, Z., Cabello-Conejo, M. I., Kidd, P., Nishizawa, N. K., Kakei, Y., et al. (2013). Xylem exudate composition and root-to-shoot nickel translocation in *Alyssum* species. *Plant Soil* 373, 59–75. doi: 10.1007/s11104-013-1782-1

Chaney, R. L., Angle, J. S., Baker, A. J. M., and Li, Y.-M. (1999). *Method for Phytomining of Nickel, Cobalt and Other Metals from Soil*. US Patent No. 5,944,872 issued Aug. 31, 1999. (Continuation-in-Part of US Patent 5,711,784 issued Jan. 27, 1998).

Chaney, R. L., Broadhurst, C. L., and Centofanti, T. (2010). "Phytoremediation of soil trace elements," in *Trace Elements in Soil,* ed P. S. Hooda (Chichester, UK: Wiley). doi: 10.1002/9781444319477.ch14

Ghaderian, S. M., Ghasemi, R., and Hajihashemi, F. (2015). Interaction of nickel and manganese in uptake, translocation and accumulation by the nickel-hyperaccumulator plant, *Alyssum bracteatum* (Brassicaceae). *Aust. J. Bot.* 63, 47–55. doi: 10.1071/bt14210

Gove, B., Hutchinson, J. J., Young, S. D., Craigon, J., and McGrath, S. P. (2002). Uptake of metals by plants sharing a rhizosphere with the hyperaccumulator *Thlaspi caerulescens*. *Int. J. Phytoremed.* 4, 267–281. doi: 10.1080/15226510208500087

Grimold, A. A., Kavanova, M., Lattanzi, F. A., and Schnyder, H. (2005). Phosphorus nutrition-mediated effects of arbuscular mycorrhiza on leaf morphology and carbon allocation in perennial ryegrass. *New Phytol.* 168, 435–444. doi: 10.1111/j.1469-8137.2005.01500.x

Jiang, C.-A., Wu, Q.-T., Goudon, R., Echevarria, G., and Morel, J.-L. (2015). Biomass and metal yield of co-cropped *Alyssum murale* and *Lupinus albus*. *Aust. J. Bot.* 63, 159–166. doi: 10.1071/bt14261

Jiang, C.-A., Wu, Q.-T., Sterckeman, T., Schwartz, C., Sirguey, C., Ouvrard, S., et al. (2010). Co-planting can phytoextract similar amounts of cadmium and zinc to mono-cropping from contaminated soils. *Ecol. Eng.* 36, 391–395. doi: 10.1016/j.ecoleng.2009.11.005

Kerkeb, J., and Krämer, U. (2003). The role of free histidine in xylem loading of nickel in *Alyssum lesbiacum* and *Brassica juncea*. *Plant Physiol.* 131, 716–724. doi: 10.1104/pp102.010686

Krämer, U. (2010). Metal hyperaccumulation in plants. *Ann. Rev. Plant Biol.* 61, 517–534. doi: 10.1146/annurev-arplant-042809-112156

Krämer, U., Grime, G. W., Smith, J. A. C., Hawes, C. R., and Baker, A. J. M. (1997). Micro-PIXE as a technique for studying nickel localization in leaves of the hyperaccumulator *Alyssum lesbiacum*. *Nuclear Inst. Methods Phys. Res. B* 130, 346–350. doi: 10.1016/S0168-583X(97)00368-6

Küpper, H., Lombi, E., Zhao, F.-J., Wieshammer, G., and McGrath, S. P. (2001). Cellular compartmentation of nickel in the hyperaccumulators *Alyssum lesbiacum, Alyssum bertolonii* and *Thlaspi goesingense*. *J. Exp. Bot.* 52, 2291–2300. doi: 10.1093/jexbot/52.365.2291

Li, Y.-M., Chaney, R., Brewer, E., Rosenberg, R., Angle, J. S., Baker, A., Reeves, R., et al. (2003b). Development of a technology for commercial phytoextraction of nickel: economic and technical considerations. *Plant Soil* 249, 107–115. doi: 10.1023/A:1022527330401

Li, Y.-M., Chaney, R. L., Brewer, E., Angle, J. S., and Nelkin, J. (2003a). Phytoextraction of nickel and cobalt by hyperaccumulator *Alyssum* species grown on nickel-contaminated soils. *Environ. Sci. Technol.* 37, 1463–1468. doi: 10.1021/es0208963

Marmiroli, M., Gonelli, C., Maestri, E., Gabrielli, R., and Marmiroli, N. (2004). Localisation of nickel and mineral nutrients Ca, K, Fe, Mg by scanning electron microscopy microanalysis in tissues of the nickel-hyperaccumulator *Alyssum montanum* L. *Plant Biosyst.* 138, 231–243. doi: 10.1080/11263500400011126

Marschner, H., and Römheld, V. (1994). Strategies of plants for acquisition of iron. *Plant Soil* 165: 261–274. doi: 10.1007/BF00008069

McNear, D. H., Chaney, R. L., and Sparks, D. L. (2010). The hyperaccumulator *Alyssum murale* uses complexation with nitrogen and oxygen donor ligands for Ni transport and storage. *Phytochemistry* 71, 188–200. doi: 10.1016/j.phytochem.2009.10.023

McNear, D. H., and Kupper, J. V. (2013). Mechanisms of trichome-specific Mn accumulation and toxicity in the Ni hyperaccumulator *Alyssum murale*. *Planta* 227, 1363–1375. doi: 10.1007/s11104-013-2003-7

McNear, D. Jr., Peltier, E., Everhart, J., Chaney, R. L., Newville, M., Rivers, M., et al. (2005). Application of quantitative fluorescence and absorption-edge computed microtomography to image metal compartmentalization in *Alyssum murale*. *Environ. Sci. Technol.* 39, 2210–2218. doi: 10.1021/es0492034

Nkrumah, P. N., Baker, A. J. M., Chaney, R. L., Erskine, P. D., Guillame, E., Morel, J.-L., et al. (2016). Current status and challenges in developing nickel phytomining: an agronomic perspective. *Environ. Pollut.* doi: 10.1007/s11104-016-2859-4

Psaras, G. K., Constantinidis, T. H., Cotsopoulos, B., and Maneta, Y. (2000). Relative abundance of nickel in the leaf epidermis of eight hyperaccumulators: evidence that the metal is excluded from both guard cells and trichomes. *Ann. Bot.* 86, 73–78. doi: 10.1006/anbo.2000.1161

Rajkumar, M., Ma, Y., and Freitas, H. (2013). Improvement of Ni phytostabilization by inoculation of Ni resistant *Bacillus megaterium* SR28C. *J. Environ. Manage.* 128, 973–980. doi: 10.1016/j.jenvman.2013.07.001

Rajkumar, M., Vara Prasad, M. N., Freitas, H., and Ae, N. (2009). Biotechnological applications of serpentine soil bacteria for phytoremediation of trace metals. *Crit. Rev. Biotechnol.* 29, 120–130. doi: 10.1080/07388550902913772

Reeves, R. J., and Adigüzel, N. (2008). The nickel hyperaccumulating plants of the serpentines of Turkey and adjacent areas: a review with new data. *Turk. J. Biol.* 32, 143–153.

Sterckeman, T., Duquene, L., Perriguey, J., and Morel, J.-L. (2005). Quantifying the effect of rhizosphere processes on the availability of soil cadmium and zinc. *Plant Soil* 276, 335–345. doi: 10.1007/s11104-005-5087-x

Tappero, R., Peltier, E., Grafe, M., Heidel, K., Ginder-Vogel, M., Livi, K. J. T., et al. (2007). Hyperaccumulator *Alyssum murale* relies on a different metal storage mechanism for cobalt than for nickel. *New Phytol.* 175, 641–654. doi: 10.1111/j.1469-8137.2007.02134.x

van der Ent, A., Baker, A. J. M., Reeves, R. D., and Chaney, R. L. (2015). Agromining: farming for metals in the future? *Environ. Sci. Technol.* 49, 4773–4780. doi: 10.1021/es506031u

van der Ent, A., Baker, A. J. M., Reeves, R. D., Pollard, A. J., and Schat, H. (2013). Hyperaccumulators of metal and metalloidtrace elements: facts and fiction. *Plant Soil* 362, 319–334. doi: 10.1007/s11104-012-1287-3

Visioli, G., D'Egidio, S., and Sanangelantoni, A. M. (2015). The bacterial rhizobiome of hyperaccumulators: future perspectives based on omics analysis and advanced microscopy. *Front. Plant Sci.* 5:752. doi: 10.3389/fpls.2014.00752

Wang, A. S., Angle, J. S., Chaney, R. L., and Delorme, T. A. (2006). Soil pH effects on uptake of Cd and Zn by *Thlaspi caerulescens*. *Plant Soil* 281, 325–337. doi: 10.1007/s11104-005-4642-9

Wang, K., Huang, H., Zhu, Z., Li, T., He, Z., Yang, X., et al. (2013). Phytoextraction of metals and rhizoremediation of PAHs in co-contaminated soil by co-planting of *Sedum alfredii* with ryegrass (*Lolium perenne)* or castor (*Ricinus communis*). *Int. J. Phytorem.* 15, 283–298. doi: 10.1080/15226514.2012.694501

Whiting, S. N., Leake, J. R., McGrath, S. P., and Baker, A. J. M. (2001a). Assessment of Zn mobilization in the rhizosphere of *Thlaspi caerulescens* by bioassay with non-accumulator plants and soil extraction. *Plant Soil* 237, 147–156. doi: 10.1023/A:1013365617841

Whiting, S. N., Leake, J. R., McGrath, S. P., and Baker, A. J. M. (2001b). Hyperaccumulation of Zn by *Thlaspi caerulescens* can ameliorate Zn toxicity in the rhizosphere of cocropped *Thlaspi arvense*. *Environ. Sci. Technol.* 35, 3237–3241. doi: 10.1021/es010644m

Wu, Q. T., Wei, Z. B., and Ouyang, Y. (2007). Phytoextraction of metal-contaminated soil by *Sedum alfredii* H: effects of chelator and co-planting. *Water Air Soil Pollut.* 180, 131–139. doi: 10.1007/s11270-006-9256-1

Zhang, F.-S., Römheld, V., and Marschner, H. (1991a). Diurnal rhythm of release of phytosiderophores and uptake rate of zinc in iron-deficient wheat. *Soil Sci. Plant Nutr.* 37, 671–678. doi: 10.1080/00380768.1991.10416935

Zhang, F.-S., Treeby, M., Römheld, V., and Marschner, H. (1991b). Mobilization of iron by phytosiderophores as affected by other nutrients. *Plant Soil* 130, 173–178. doi: 10.1007/BF00011872

Zhang, L., Angle, J. S., and Chaney, R. L. (2007). Do high-nickel leaves shed by the Ni-hyperaccumulator Alyssum murale inhibit seed germination of competing plants? *New Phytol.* 173, 509–516. doi: 10.1111/j.1469-8137.2006.01952.x

Zhao, F. J., Harmon, R. E., and McLaughlin, M. J. (2001). Root exudates of the hyperaccumulator *Thlaspi caerulescens* do not enhance metal mobilization. *New Phytol.* 151, 613–620. doi: 10.1046/j.0028-646x.2001.00213.x

Zuo, Y., and Zhang, F. (2011). Soil and crop management strategies to prevent iron deficiency in crops. *Plant Soil* 339, 83–95. doi: 10.1007/s11104-010-0566-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Broadhurst and Chaney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 March 2016
doi: 10.3389/fpls.2016.00259

De novo Transcriptome Analysis of *Sinapis alba* in Revealing the Glucosinolate and Phytochelatin Pathways

*Xiaohui Zhang, Tongjin Liu, Mengmeng Duan, Jiangping Song and Xixiang Li **

Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China

Sinapis alba is an important condiment crop and can also be used as a phytoremediation plant. Though it has important economic and agronomic values, sequence data, and the genetic tools are still rare in this plant. In the present study, a *de novo* transcriptome based on the transcriptions of leaves, stems, and roots was assembled for *S. alba* for the first time. The transcriptome contains 47,972 unigenes with a mean length of 1185 nt and an N50 of 1672 nt. Among these unigenes, 46,535 (97%) unigenes were annotated by at least one of the following databases: NCBI non-redundant (Nr), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, Gene Ontology (GO), and Clusters of Orthologous Groups of proteins (COGs). The tissue expression pattern profiles revealed that 3489, 1361, and 8482 unigenes were predominantly expressed in the leaves, stems, and roots of *S. alba*, respectively. Genes predominantly expressed in the leaf were enriched in photosynthesis- and carbon fixation-related pathways. Genes predominantly expressed in the stem were enriched in not only pathways related to sugar, ether lipid, and amino acid metabolisms but also plant hormone signal transduction and circadian rhythm pathways, while the root-dominant genes were enriched in pathways related to lignin and cellulose syntheses, involved in plant-pathogen interactions, and potentially responsible for heavy metal chelating, and detoxification. Based on this transcriptome, 14,727 simple sequence repeats (SSRs) were identified, and 12,830 pairs of primers were developed for 2522 SSR-containing unigenes. Additionally, the glucosinolate (GSL) and phytochelatin metabolic pathways, which give the characteristic flavor and the heavy metal tolerance of this plant, were intensively analyzed. The genes of aliphatic GSLs pathway were predominantly expressed in roots. The absence of aliphatic GSLs in leaf tissues was due to the shutdown of *BCAT4*, *MAM1*, and *CYP79F1* expressions. Glutathione was extensively converted into phytochelatin in roots, but it was actively converted to the oxidized form in leaves, indicating the different mechanisms in the two tissues. This transcriptome will not only benefit basic research and molecular breeding of *S. alba* but also be useful for the molecular-assisted transfer of beneficial traits to other crops.

Keywords: transcriptome, *Sinapis alba*, glucosinolate, phytochelatin, SSR marker, deep sequencing

OPEN ACCESS

Edited by:
Naser A. Anjum, University of Aveiro, Portugal

Reviewed by:
Lijun Chai, Huazhong Agricultural University, China
Yuyang Zhang, Huazhong Agricultural University, China

****Correspondence:***
Xixiang Li
lixixiang@caas.cn

Specialty section:
This article was submitted to Crop Science and Horticulture, a section of the journal Frontiers in Plant Science

Received: *17 November 2015*
Accepted: *17 February 2016*
Published: *04 March 2016*

Citation:
Zhang X, Liu T, Duan M, Song J and Li X (2016) De novo Transcriptome Analysis of Sinapis alba in Revealing the Glucosinolate and Phytochelatin Pathways. Front. Plant Sci. 7:259. doi: 10.3389/fpls.2016.00259

INTRODUCTION

Sinapis alba, known as yellow mustard or white mustard, is an important cruciferous crop widely used as food condiments in the world (Hemingway, 1995). It has many desirable agronomic traits, such as tolerance or resistance to drought, disease, pests, and pod-shattering (Thompson, 1963; Bodnaryk and Lamb, 1991; Brown et al., 1997; Lee et al., 2014), making it an attractive resource for oil crop breeding (Tian et al., 2014). The genetically close relationship and the ease of forming hybrids between *S. alba* and *Brassica* plants make it a potential donor of resistant and other agronomic traits to *Brassica napus* and other *Brassica* crops (Brown et al., 1997; Jiang et al., 2013; Lee et al., 2014). Recently, the discovery of anti-bacterial, antioxidant, and anticancer agents in the seed extract of *S. alba* increased the interest in this plant and expanded its application beyond spices (Zielniok et al., 2015).

The spicy "heat" sensation of the *S. alba* seed powder is caused by the hydrolysis products of glucosinolates (GSLs; Hemingway, 1995; Javidfar and Cheng, 2013). The anti-bacterial and carcinogenesis-inhibiting activities of this plant are also attributed to the GSLs and their derivatives (Peng et al., 2014). Some GSL-hydrolyzed products, such as 4-methylsulfanyl-3-butenyl isothiocyanate, have been experimentally proven of their potential chemo- and cancer-prevention abilities (Abdull Razis et al., 2012). These applications exploit the advantages of GSLs. However, a high GSL content is a defective trait when attempting to use *S. alba* as an oil seed crop. Thus, GSL is a primary trait for this plant, and modulating the GSL type and its content for different application goals is important for breeding processes. GSLs belong to a type of nitrogen- and sulfur-containing plant secondary metabolite that widely exist in the order Brassicales (Fahey et al., 2001; Grubb and Abel, 2006). The GSL metabolic pathway has been extensively investigated in *Arabidopsis* and has been well studied in *B. rapa*, broccoli, radish, etc. by genome wide homologous analysis (Wittstock and Halkier, 2002; Zang et al., 2009; Wang et al., 2011; Liu et al., 2014; Pino Del Carpio et al., 2014; Wiesner et al., 2014; Mitsui et al., 2015). In *S. alba*, many GSLs have been identified (Agerbirk et al., 2008; Popova and Morra, 2014; Vastenhout et al., 2014), and many studies on the functions of GSLs have been reported (Abdull Razis et al., 2012; Peng et al., 2014). Furthermore, a QTL mapping of GSL contents has been carried out (Javidfar and Cheng, 2013). However, knowledge on the metabolic pathway of this plant is still limited.

In addition to the applications mentioned above, *S. alba* can also be used as a phytoremediation plant due to its outstanding ability to absorb cadmium (Cd) and its high biomass productivity (Plociniczak et al., 2013). Cd tolerance is potentially related to a series of biological characteristics and physiological processes, such as barriers by cell walls or mycorrhizas, reduced uptake, or efflux pumping by plasma membranes, chelation by phytochelatins, or metallothioneins, and compartmentation to vacuoles (Hall, 2002). Phytochelatin is believed to be one of the most important factors mediating Cd tolerance by chelating the heavy metal and facilitating its transport to storage locations in order to avoid cell toxicity (Cobbett, 2000; Mendoza-Cázatl et al., 2011). Phytochelatin is a stretch of $(\gamma\text{-Glu-Cys})_n$-Gly peptides produced in plants not via translation but by a biosynthesis process catalyzed by γ-glutamylcysteine dipeptidyl transpeptidase (phytochelatin synthase), with glutathione as the substrate (Grill et al., 1989; Cobbett, 2000).

Despite, its agricultural importance and prospective applications, genetic research on *S. alba* is far behind that of other cruciferous crops such as *B. napus* and *B. rapa*, and publicly released sequence data are rare. Though, Illumina sequencing technology has been rapidly developing and has been successfully used for many years, to the best of our knowledge, only one RNA-Seq project on *S. alba* has been published. That study profiled the differential expressions of *S. alba* leaves between drought and water recovery conditions using a transcript-end sequencing strategy in which *de novo* transcriptome assembly was not applicable (Dong et al., 2012). In the present study, a *de novo* transcriptome assembly was carried out by Illumina sequencing of mRNAs from root, stem, and leaf tissues. The transcriptome was annotated, and SSRs were identified to facilitate the application. Additionally, the pathways of GSLs and phytochelatins were analyzed.

MATERIALS AND METHODS

Plant Materials and RNA Extraction

For transcriptome sequencing, *S. alba* (ZYZ-1553) was sown in plastic pots (20 cm wide × 20 cm deep) filled with a mixture of peat soil (peat:moss:perlite:vermiculite soil = 3:2:1:1). Each pot contained one plant and was placed in a plastic tunnel located at the experimental farm of the Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China. Plants were regularly watered and fertilized. The sowing date was September 25th, and the sampling date was November 10th 2013. The leaf, stem, and root tissues were sampled from three individual plants at vegetative developmental stage (**Figure 1**) and then snap-frozen in liquid nitrogen and kept at −80°C for further use. Total RNA was extracted using the TRIzol reagent (Invitrogen, USA). DNase (Promega, USA) was used to remove potential DNA contamination.

For Quantitative PCR (qPCR) analysis to backup the transcriptome profiling, the plants were grown in a greenhouse with similar condition in the winter of 2015. Total RNAs were isolated from three independent plants with the same method.

cDNA Library Construction and Illumina Sequencing

Total RNA (10 μg) was subjected to poly-A selection, fragmentation, random priming, and first and second strand cDNA synthesis with the Illumina Gene Expression Sample Prep kit (CA, USA). The cDNA fragments were subjected to an end repair process and then ligated to adapters. The products were enriched with PCR, and the fragments harboring 330-bp inserts were purified with 6% TBE PAGE gel electrophoresis. After denaturation, the single-chain fragments were fixed onto

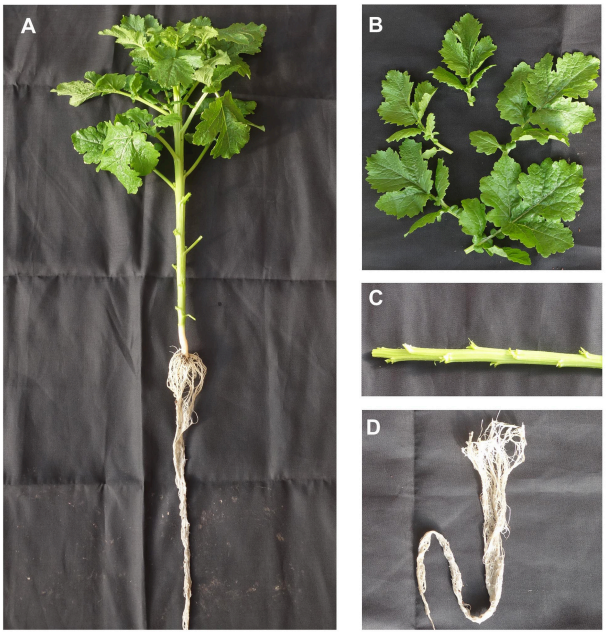


FIGURE 1 | Plants and sampling. (A) Whole plant for sampling. **(B)** Leaf sample. **(C)** Stem sample. **(D)** Root sample.

the Solexa Sequencing Chip (Flowcell) and consequently grown into single-molecule cluster sequencing templates through in situ amplification on the Illumina Cluster Station. Double-end pyrosequencing was performed on the Illumina Genome Analyzer platform with read lengths of 100 bp for each end.

Assembly

Raw reads were first subjected to purification by removal of adaptors and low quality reads. The clean reads of leaf, stem, and root tissues were separately subjected to transcriptome *de novo* assembly using the short-read assembling program Trinity (Grabherr et al., 2011). The longest assembled sequences were termed as contigs. The paired-end reads were then mapped back to the contigs. Sequences without gaps and could not be extended at either end were defined as transcripts. The transcripts were then assembled into unigenes by filtering out redundant sequences and further assembled using TGI Clustering Tool (TGICL; Pertea et al., 2003). The unigenes from the three samples were clustered again; the longest sequences from the three data sets were adopted to form a single set of non-redundant unigenes. The unigenes were divided into two classes: Clusters including several unigenes with more than 70% similarity were prefixed with CL and suffixed with an ID number, and singletons that could not cluster to other genes were prefixed Unigene and followed with an ID number suffix.

Annotation

We searched all unigene sequences against protein databases (Nr, Swiss-Prot, KEGG, and COG) using BLASTX (e- $< 10^{-5}$). Protein function information was predicted from annotation of the most similar proteins in those databases. Proteins with the highest ranks in the BLAST results were obtained to determine the coding region sequences (CDS) of the unigenes, after which CDS were translated into amino sequences using the standard codon table. Unigenes that could not be aligned to any database were scanned by ESTScan (Iseli et al., 1999), producing the nucleotide sequence (5′–3′) direction and amino sequence of the predicted coding region.

Expression Levels

Unigene expression levels were calculated using the reads per kilobase per million (RPKM) method (Mortazavi et al., 2008), and the formula used is RPKM = (1,000,000 * C) / (N * L * 1000), where RPKM(A) is the expression of gene A, C is the number of reads that uniquely align to gene A, N is the total number of reads that uniquely align to all genes, and L is the length of gene A. Statistical comparisons between two samples were performed using the IDEG6 software (Romualdi et al., 2003). The general Chi squared method was used, and the false discovery rate (FDR) was applied to determine the threshold of the *Q*-value. Unigenes were considered differentially expressed (DEG) when the RPKM between two samples displayed a more than two-fold change, with an FDR $< 10^{-3}$.

GO and KEGG Enrichments

The differentially expressed genes (DEGs) were mapped to GO terms and KEGG pathways and then subjected to an enrichment analysis using a hypergeometric test to find over-represented GO terms and KEGG pathways. The algorithm used is described as follows:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

where N is the number of all genes with GO or KEGG annotation, n is the number of DEGs in N, M is the number of all genes that are annotated to certain GO terms or KEGG pathways, and m is the number of DEGs in M. The calculated *p*-value goes through Bonferroni Correction (Abdi, 2007), taking a corrected $p \leq 0.05$ as the threshold.

Simple Sequence Repeat (SSR) Mining

The MIcroSAtellite identification tool MISA (http://pgrc.ipk-gatersleben.de/misa/) was used to identify and localize SSRs in unigenes longer than 1 kb. The SSR-containing sequences were extracted with a 300-bp (if <300 bp, extracted from the end) fragment upstream and downstream of the SSR region sequence to facilitate primer design. SSR primers were design by Primer 3.0 (Untergasser et al., 2012).

Identification of Glucosinolate and Phytochelatin Pathways

For the glucosinolate pathway, candidate genes were first identified by BLASTN ($e < 10^{-100}$) search of *S. alba* unigenes using *Arabidopsis* glucosinolate biosynthesis and transcription factors as baits. Sequences representing the complete set of

glucosinolate biosynthetic and regulator genes in *A. thaliana* were acquired from the TAIR database (www.arabidopsis.org). Unigenes annotated ($e < 10^{-10}$) to the KEGG glucosinolate biosynthesis reference pathway (map00966) were identified. The candidate genes were finally identified from these two selections by a manual check.

For glutathione pathways, the genes were identified from KEGG annotation to the glutathione metabolism reference pathway (map00480) with $e < 10^{-100}$. Only the two directly related cycling pathways were adopted in this study. The phytochelatin synthase was identified from KO (K05941) annotation ($e = 0$).

Quantitative PCR (qPCR)

Total RNAs (800 ng) were synthesized to first-strand cDNAs templates using *EasyScript* One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen, Beijing, China). Experiments were performed on a *Mastercycler ep realplex* Real-Time PCR System (Eppendorf, Germany) using *TransStar* Green qPCR SuperMix (TransGen). The genes and primers were listed in Table S1. The reaction volume was 25 μL, including 0.5 μL of 10 mM Forward and Reverse primer, respectively, 12.5 μL of 2 × *TransStart* Green qPCR SuperMix, 2.0 μL of the cDNA templates, 0.5 μL of Passive Reference Dye I, and 9 μL of ddH_2O. The thermal cycling profile was: 95°C for 30 s; 40 cycles of 95°C for 10 s, 58°C for 15 s, 72°C for 10 s; then 95°C for 15 s, 60°C for 1 min, ramping to 95°C for 15 s. Three independent biological and two technical replicates were performed. *GAPDH* was used as an internal control. The relative expression levels were estimated by the $2^{-\Delta\Delta CT}$ method.

RESULTS AND DISCUSSION

Transcriptome Sequencing and Assembly

To construct a *de novo* transcriptome database, three mRNA libraries were generated from the root, stem and leaf tissues of *S. alba* by Illumina sequencing. ∼26.3, 27.9, and 26.4 million paired-end reads (100 bp read length) containing 5.26, 5.59, and 5.28 gigabase pairs of nucleotides were generated for the three samples, respectively, (**Table 1**). After filtering out low quality reads, ∼24.2 million clean paired-end reads containing ∼4.8 gigabase of clean nucleotides were obtained for each tissue. The overall GC percentages were 46.6–48.6% in these tissues. The reads of the three tissues were first assembled separately into three distinct sets of contigs and unigenes, which were consequently combined and further assembled into a set of 47,972 non-redundant unigenes, with a mean length of 1185 nt and a N50 length of 1672 nt (**Table 1**). The length distributions of the unigenes are shown in **Figure 2**, indicating a high quality reference transcriptome for use in future studies. All of the unigene sequences are provided in Supplementary File 1.

Functional Annotation

We screened the unigene sequences against the NCBI non-redundant (Nr), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, Gene Ontology (GO), and Clusters of Orthologous Groups of proteins (COGs) protein databases using BLASTX ($e < 10^{-5}$). Unigenes were also searched against the NCBI non-redundantnucleotide sequence (Nt) database using BLASTN ($e < 10^{-5}$). Protein function was predicted from the annotations of the most similar proteins in those databases. In total, 46,535 (97.0%) of the 47,972 unigenes were annotated by at least one of these databases (**Table 1**). Amongst, 44,485 (92.7%) unigenes were annotated by Nr. As shown in **Figure 3**, more than 87.8% of the unigenes were annotated with an $e < 10^{-30}$, and more than 70.9% of the unigenes contains more than 80% similarity to the reference genes in the database, indicating that the annotations are reliable. ∼94.2% of the unigenes were annotated to cruciferous plants (**Figure 3C**).

TABLE 1 | Summary of the assembly and annotation of the *Sinapis alba* transcriptome.

	Root	Stem	Leaf
Total reads	26,300,732	27,933,350	26,397,896
Nucleotides (nt)	5,260,146,400	5,586,670,000	5,279,579,200
Clean reads	24,210,304	24,177,099	24,210,304
Clean nucleotides (nt)	4,842,060,800	4,835,419,800	4,842,060,800
GC percentage (%)	47.85	46.63	48.64
Unigene number	53,306	49,153	45,745
Mean length of unigene (nt)	900	862	760
Combined non-redundant unigene		47,972	
Total length (nt)		56,824,691	
Mean length (nt)		1185	
N50 (nt)		1672	
Nr		44,485 (92.73%)	
Nt		46,140 (96.18%)	
GO		41,796 (87.13%)	
COG		18,906 (39.41%)	
Swiss-Prot		31,614 (65.9%)	
KEGG		27,323 (56.96%)	
All annotated		46,535 (97%)	

COGs annotation indicated that 18,906 (39.4%) unigenes were assigned to one or more COG functional classes. The most abundant class was "general function prediction only," including 6946 (36.7% of the annotated COGs) unigenes, followed by the classes "transcription" (3866; 20.5%) and "replication, recombination and repair" (3146; 16.6%; **Figure 4A**). Functions of 41,796 (87.1%) unigenes were further classified by Gene Ontology (GO) analysis. The largest GO terms found in the "biological process" ontology were "cellular process" and "metabolic process," comprising 71.1 and 68.7% of the GO-termed unigenes, respectively. In the "cellular component" and "molecular function" ontologies, the top terms were "cell (or cell part)" and "binding," which are 90.9 and 50.9% of the total unigenes annotated by GO, respectively, (**Figure 4B**). KEGG metabolic pathway analysis revealed that 27,323 (56.96%) unigenes could be assigned to 128 pathways (level 3). The most abundant pathways are metabolic, secondary metabolite biosynthesis, and plant hormone signal transduction, comprising 5891 (21.56%), 2868 (10.5%), and 1853 (6.78%) unigenes,

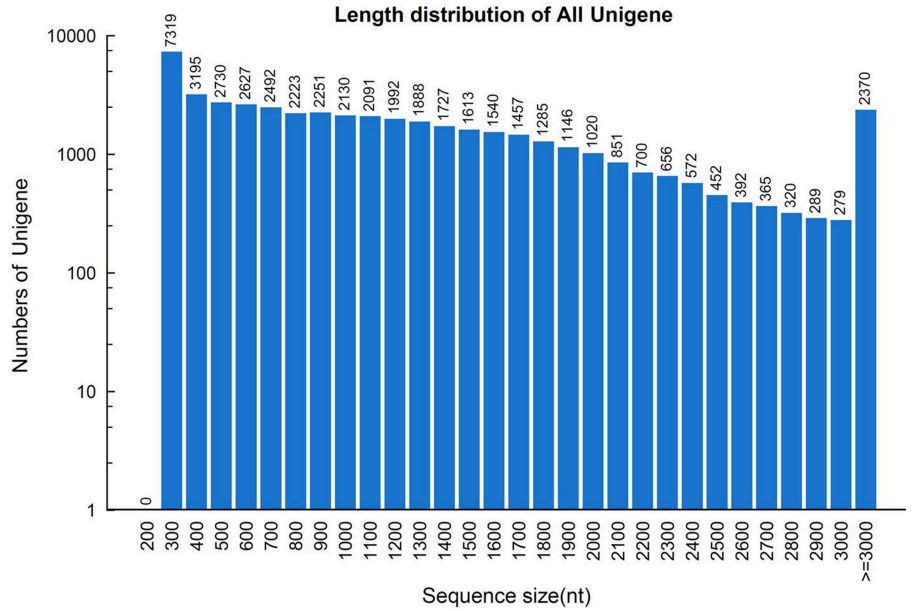


FIGURE 2 | Length distribution of the unigene assembly in the *Sinapis alba* transcriptome.

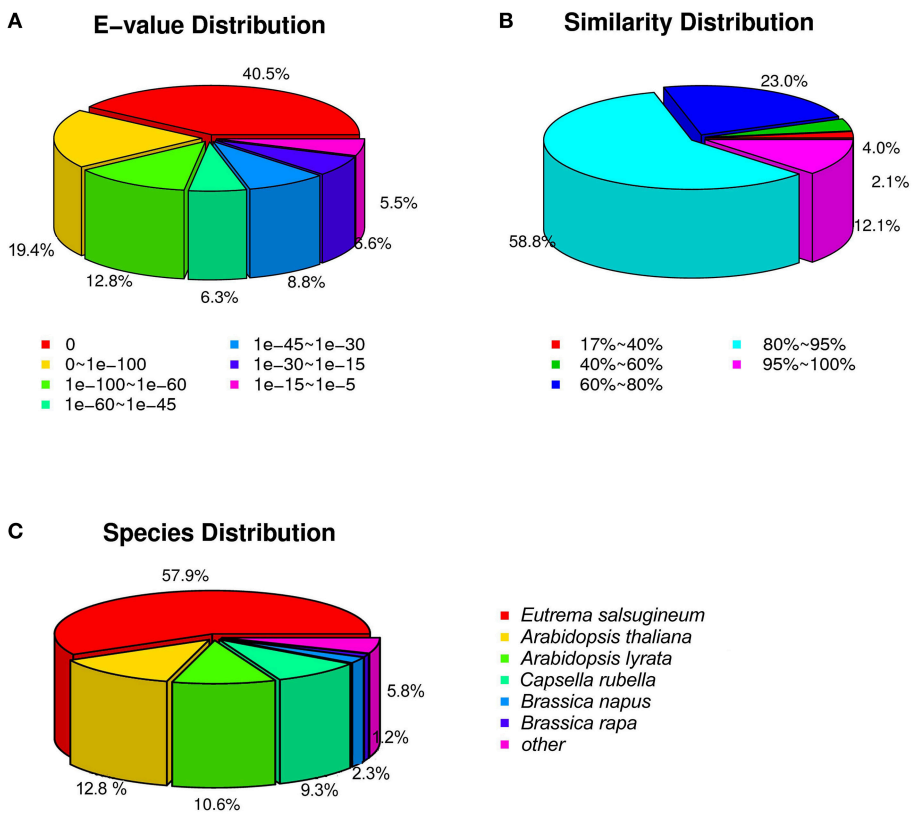


FIGURE 3 | The distribution of unigenes annotated to the NCBI non-redundant nucleotide sequence (Nt) database. (A) *E*-value distribution. **(B)** Similarity distribution. **(C)** Species distribution.

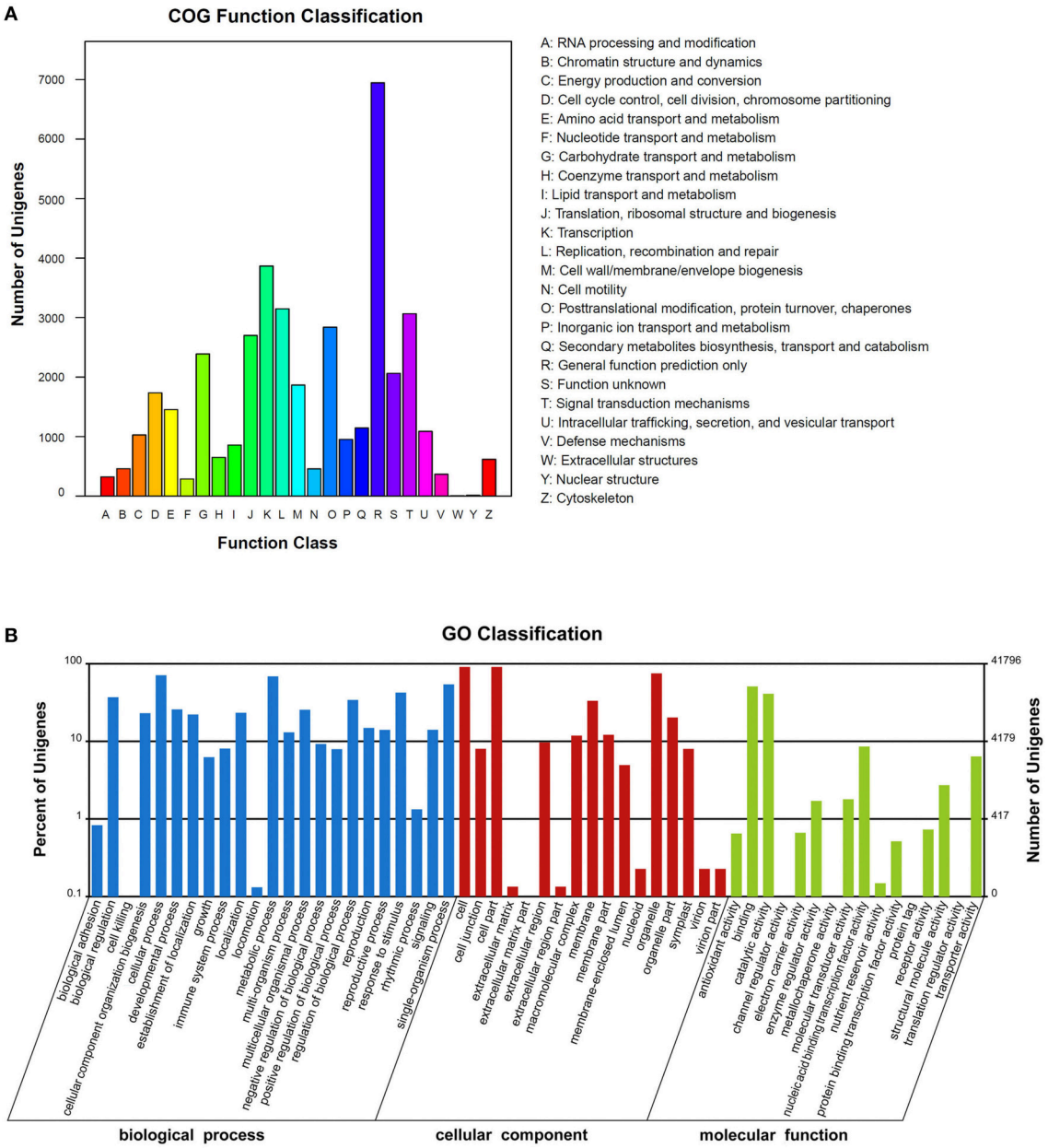


FIGURE 4 | The clusters of unigenes of the *Sinapis alba* transcriptome. (A) Clusters of Orthologous Groups of proteins (COGs) annotation. **(B)** Gene Ontology (GO) terms.

respectively, (Supplementary File 2). All of the above annotations to each unigene are integrated in Table S2.

The orientation and coding sequence (CDS) of 43,953 unigenes were determined by BLASTX ($e < 10^{-5}$) to Nr, Swiss-Prot, KEGG, and COG databases. Those unigenes that had no blast hit to any database were analyzed by ESTScan, in which 502 additional unigenes were assigned an orientation and a CDS. The encoding proteins were deduced from the CDSs using the standard codons, and the protein sequences are shown in Supplementary File 3.

Tissue Patterns

To profile the expressional tissue patterns, we first aligned the reads back to the unigenes; the reads aligned to mono sites were counted for expression calculations. 1644, 286, and 338 unigenes were specifically expressed in roots, stems and leaves, respectively, (**Figure 5A**). The fact that more genes were specifically expressed in roots than in stems and leaves indicated, that the root system faced more complications and performed many specific functions due to rhizosphere microbe. The root-specific genes included a large fraction of transcription

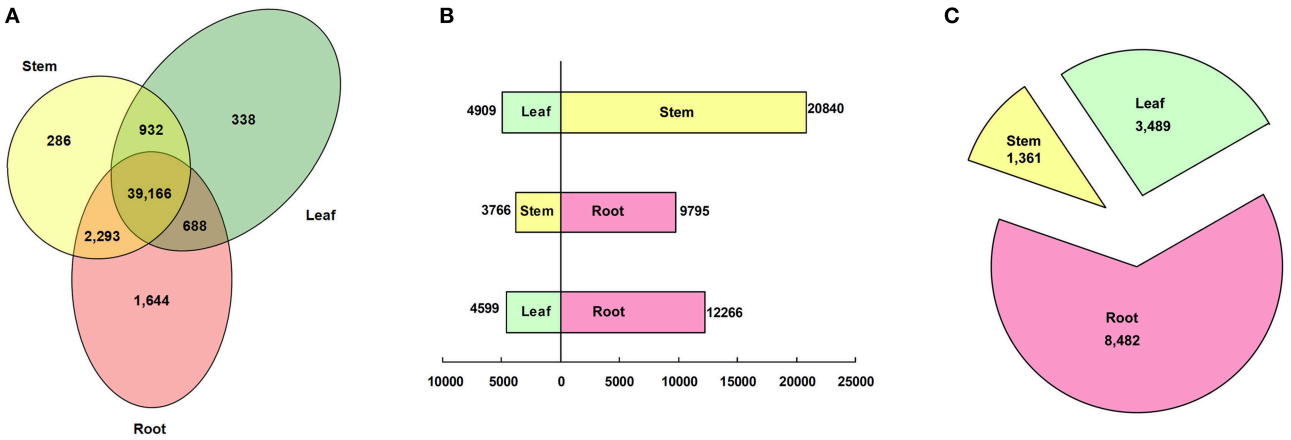


FIGURE 5 | Classification of tissue patterns. (A) Venn diagram of tissue-specific gene expression. **(B)** Histogram of differentially expressed genes between tissues. **(C)** Pie chart of genes predominantly expressed in each tissue.

factors (72, 4.38%) and genes related to phytohormone (33, 2.01%) and materials transport (102, 6.20%; Table S3). The transcription factors were composed of 33 MYBs, 22 MADS-boxes (including 10 AP2 and 7 EREBP-like), 7 WRKYs, 4 TGA, 3 homeobox-leucine zipper proteins and 2 MYC2. The root-specific phytohormone genes were related to the metabolism and signal transduction of gibberellins (GA), jasmonate (JA), abscisic acid (ABA), brassinosteroid (BR), and cytokinin. The majority of root-specific transporters were comprised of 55 major facilitator super family (MFS) transporters which facilitate the transport of glucose, sugar, amino acid, peptide, histidine, zinc, nitrate/nitrite, inorganic phosphate, and organic cation, as well as serving as sodium/hydrogen exchangers and iron-regulated transporters. Other transporters included cation transport ATPase, multidrug resistance protein, ABC-type multidrug transport system, Ca^{2+}/H^{+} antiporter, K^{+} transporter, magnesium transporter, ammonium transporter, amino acid transporter, copper chaperone, and vacuolar protein sorting-associated protein.

Using a threshold of RPKM $>$ 2 and FDR $<$ 0.001, 25,749 (53.68%), 13,561 (28.27%), and 16,865 (35.16%) unigenes were differentially expressed between the leaf and the stem, the stem and the root, and the leaf and the root, respectively, (**Figure 5B**). By using a hypergeometric distribution analysis, the root-stem DEGs were enriched in GO terms related to the vacuole, the cell wall, and the ER body. The stem-leaf DEGs were enriched in terms related to the plastid, chloroplast, and Golgi apparatus. The GO terms of the photosystem, plastid part, apoplast, and those integral to the membrane were significant differentially expressed in the root, stem, and leaf (Table S4). Via pathway enrichment, the inter-tissue DEGs were enriched in 50 pathways (Table S5). Aside from the "endocytosis," "regulation of autophagy," "SNARE interactions in vesicular transport," "ABC transporters," "plant-pathogen interaction," "Circadian rhythm—plant," and "plant hormone signal transduction" enriched in the three comparisons. All of the other 43 pathways belonged to the metabolism pathway, including "energy metabolism," "carbohydrate metabolism," "lipid metabolism," "glycan biosynthesis and metabolism," "metabolism of terpenoids and polyketides," and "biosynthesis of other secondary metabolites." The global maps of "metabolic pathways" and "biosynthesis of secondary metabolites," "ether lipid metabolism," "porphyrin and chlorophyll metabolism," some pathways of "energy metabolism," and "carbohydrate metabolism" were enriched in any two of the three tissue comparisons. The biosynthesis of secondary metabolites, such as phenylalanine, tryptophan, stilbenoid, diarylheptanoid and gingerol, flavones, flavonoid, indole alkaloid, and phenylpropanoid were enriched in the root-stem and the root-leaf comparisons. The biosynthesis of glucosinolate, isoflavonoid, isoquinoline and benzoxazinoid, cyanoamino acid, glutathione, and seven types of terpenoids and polyketides were enriched in the root-stem comparison. The "amino sugar and nucleotide sugar metabolism," "ascorbate and aldarate metabolism," "fructose and mannose metabolism," and "pentose phosphate pathway" pathways were significantly enriched in the stem-leaf comparison.

Further, analysis identified that 3489, 1361, and 8482 unigenes were predominantly expressed (more than a two-fold up-expression than the other two samples, FDR $<$ 0.001) in the leaf, stem, and root, respectively, (**Figure 5C**). Via a hypergeometric test, the leaf-dominant genes were enriched in pathways related to photosynthesis and carbon fixation (**Figure 6A**). This finding is consistent with the fact that the leaf is the main photosynthesis organ. Genes predominantly expressed in the stem were not only enriched in pathways related to sugar, ether lipid, and amino acid metabolism but also in plant hormone signal transduction and circadian rhythm pathways (**Figure 6B**). The root dominant genes enriched in flavonoid, phenylpropanoid, and terpenoid biosynthesis pathways, plant hormone signal transduction, plant-pathogen interaction, regulation of autophagy and ABC transporters and glutathione metabolism (**Figure 6C**).The enrichment of the flavonoid and phenylpropanoid biosynthesis pathways indicates, that roots were actively synthesizing lignin and cellulose for

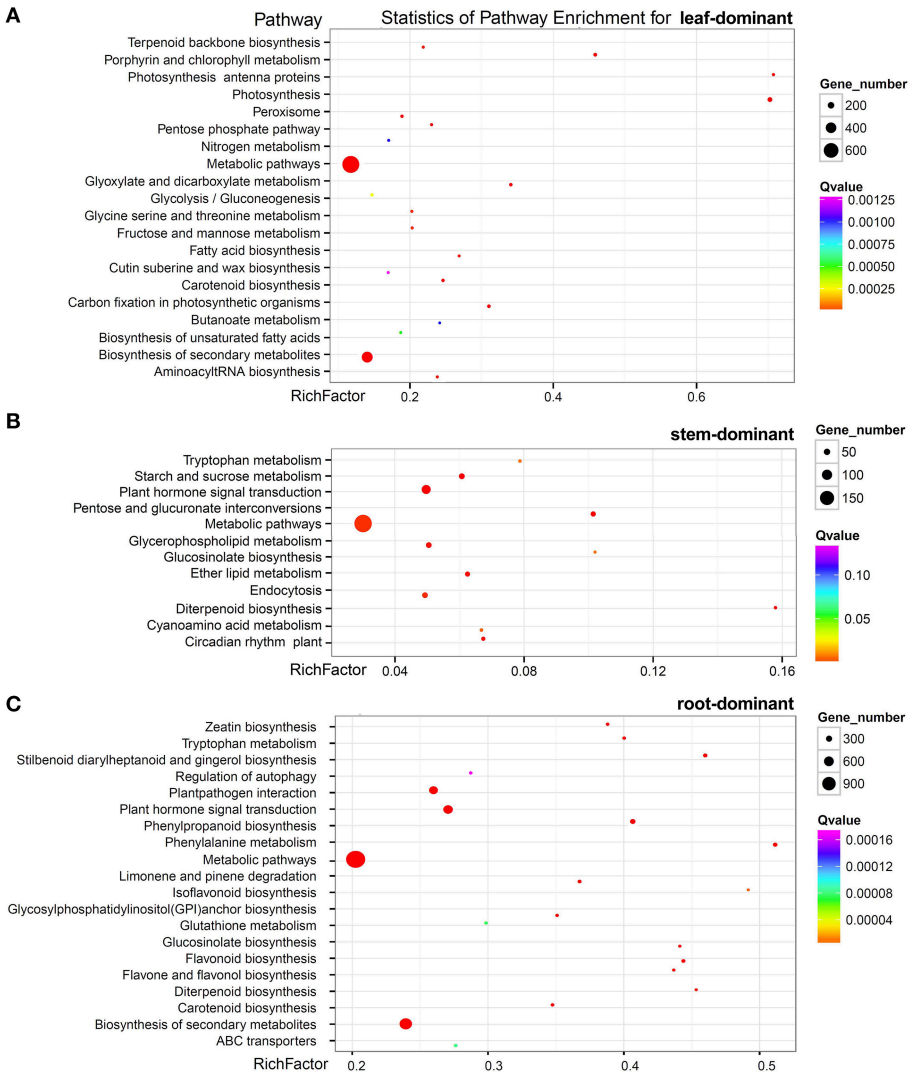


FIGURE 6 | Pathway enrichment of genes predominantly expressed in the leaf (A) the stem (B) and the root (C).

rapid cell growth. The enrichment of plant-pathogen interaction pathways indicates that roots faced soil borne diseases. The enrichment of glutathione metabolism genes is an interesting phenomenon because glutathione-generated phytochelatin plays an important role in Cd absorption (Grill et al., 1989; Cobbett, 2000). The enrichment of autophagy and ABC transporters is consistent with mineral nutrient uptake, which is the primary function of roots. The ABC transporters are also responsible for transporting phytochelatin-chelated heavy metals to the vacuolar for storage and detoxification (Mendoza-Cázatl et al., 2011).

Simple Sequence Repeat (SSR)

Molecular marker is important genetic tool but is still considered undeveloped for *S. alba*. Though many intron length polymorphism markers have been developed for this plant (Javidfar and Cheng, 2013), SSR is still a useful molecular marker, especially for this less-studied crop. In this study, a total of 14,727 SSRs was identified from 56,824,691 nt transcriptome sequences; nearly every 4953 nt contained one SSR. Among them, 11,473 (23.9%) of the 47,972 total unigenes contained at least one SSR, of which 2539 unigenes contained more than one SSR. The tri-, di-, and mono-nucleotide repeats were the main types, with values of 6809 (59.3%), 5320 (46.4%), and 2142 (18.7%) SSRs, respectively, (**Table 2**). The A/T, AG/CT, and AAG/CTT were the dominant types of the mono-, di-, and tri-nucleotide repeats (**Figure 7**).

To facilitate applications, 12,830 pairs of primers were developed for 2522 SSR-containing unigenes (Table S6). Twenty (0.16%), 2579 (20.1%), 9873 (77%), 84 (0.65%), 119 (0.93%), and 115 (1.21%) pairs of primers were designed for detecting fragments harboring mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide repeats. To the best of our knowledge, this is the first large collection of SSR markers for this plant.

TABLE 2 | Statistics of simple sequence repeat types.

Number of repeats	Mono-nucleotide repeats	Di-nucleotide repeats	Tri-nucleotide repeats	Tetra-nucleotide repeat	Penta-nucleotide repeats	Hexa-nucleotide repeats
4	0	0	0	0	175	133
5	0	0	4299	104	17	0
6	0	2104	1755	27	0	0
7	0	1302	682	0	0	0
8	0	891	71	0	0	0
9	0	530	1	0	0	0
10	0	335	1	0	0	0
11	0	145	0	0	0	0
12	669	13	0	0	0	0
13	467	0	0	0	0	0
14	336	0	0	0	0	0
15	188	0	0	0	0	0
16	152	0	0	0	0	0
17	93	0	0	0	0	0
18	55	0	0	0	0	0
19	44	0	0	0	0	0
20	13	0	0	0	0	0
21	37	0	0	0	0	0
22	45	0	0	0	0	0
23	36	0	0	0	0	0
24	7	0	0	0	0	0
Subtotal	2142	5320	6809	131	192	133

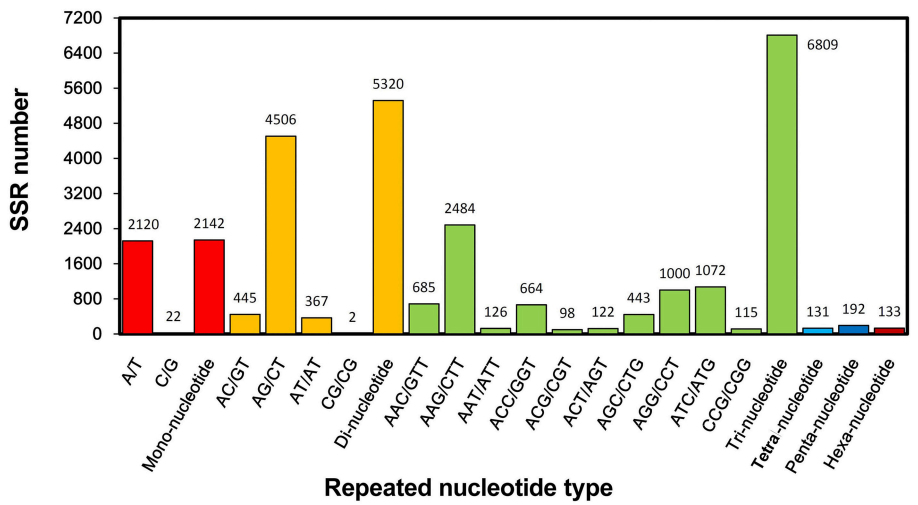


FIGURE 7 | Histogram of the statistics of simple sequence repeat (SSR) types.

Glucosinolate Metabolic Pathway in *S. alba*

Glucosinolate (GSL) is an important metabolite that confers the special pungent properties of *S. alba* seeds (Hemingway, 1995; Javidfar and Cheng, 2013). Most of the glucosinolate in *S. alba* is 4-hydroxybenzyl, 3-indolylmethyl, 4-hydroxy-3-indolylmethyl, and 2-hydroxy-3-butenyl GSL (Javidfar and Cheng, 2013). Therefore, all three GSL metabolic pathways, i.e., the aliphatic, indolic, and aromatic GSL pathways, must exist in *S. alba*. To investigate the molecular basis for the GSL biosynthesis in this plant, the transcripts of the GSL pathways were identified in the *S. alba* transcriptome by KEGG annotation and by BLAST search for homologous genes of the *Arabidopsis* GSL pathway. A total of 71 transcripts were identified as candidate genes for 32 enzymes of GSL biosynthesis and degradation pathways (Table S7). A

deduced GSL metabolic pathway map was constructed for *S. alba* (**Figure 8A**). The pathway is comprised of four stages including side chain elongation, core structure synthesis, side chain modification, and degradation of GSLs. The side chain elongation process was investigated thoroughly only for methionine in the aliphatic GSL pathway. The methionine was firstly deaminated to form 2-keto acid by a branched-chain amino acid aminotransferase (BCAT4) and then entered two cycles of three successive transformations: (1) condensation with acetyl-CoA by methylthioalkylmalate synthase 1(MAM1), (2) isomerization by isopropylmalate isomerase large subunit 1 (IPMI LSU1) and isopropylmalate isomerase small subunit 2 (IPMI SSU2), and (3) oxidative decarboxylation by isopropylmalate dehydrogenase 1 (IPMDH1; Sønderby et al., 2010; Wang et al., 2011). After a transamination reaction catalyzed by BCAT3, two carbons were added to the side chain of methionine. Five of these six enzymes (except BCAT3) displayed significantly up-regulated expression in the root compared to the stem and the leaf (**Figure 8B,** Table S7). In particular, *BCAT4* was not expressed and *MAM1* was barely detectable in the leaf, indicating that the side chain elongation pathway is blocked in this organ. The side-chain elongated methionine was then subjected to core structure synthesis. Twenty-two genes that encoded 10 enzymes catalyzing the seven steps in this process were expressed (**Figure 8B,** Table S7). First, two cytochrome P450s (CYP79F1 and CYP83A1) successively converted the side-chain elongated methionine to nitrile oxide. Then, the molecules were conjugated to a glutathione (GSH) by two glutathione S-transferases (GSTF11 and GSTU20). *CYP79F1*, *GSTF11*, and *GSTU20* were highly expressed in the root, moderately expressed in the stem and only trace amounts were expressed in the leaf. This result, again, implies that the aliphatic GSLs are predominantly produced in the root but minimally synthesized in the stem and leaf of *S. alba*. This finding is consistent with a previous report that aliphatic GSLs were detected significantly in the root but minimally detected (if any) in the leaf (Agerbirk et al., 2008). GSH conjugates were deglutamylated by gamma-glutamyl peptidase 1 (GGP1) to form S-alkyl-thiohydroximate and then catalyzed by C-S lyase SUPERROOT1 (SUR1) to form thiohydroximic acids and consequently S-glucosylated by UDP-glucosyltransferase (UGT74B1&C1) to generate desulfoglucosinolates, which were finally catalyzed by desulfoglucosinolate sulfotransferases (ST5b&c) to generate the core structure of glucosinolate. The genes for these four metabolic steps were expressed in the three organs with no significant difference, indicating these reactions were not rate-limiting nodes in this pathway.

For the indolic and aromatic pathways, the side-chain elongation pathway did not exist and was not detected in this plant. The core structure synthesis processes were similar to that of the aliphatic pathway, with the only difference being that the members from a different subfamily of P450 (CYP79A2, B2, B3; CYP83B1), GST (GST9&10), and ST (ST5a) were present in the indolic (Wiesner et al., 2014) and aromatic pathways (**Figure 8**, Table S7). Interestingly, despite the fact that *CYP79B2&3* and *CYP83B1* were still expressed significantly higher in the root, some genes including *CYP79A2* and *ST5a* were expressed the highest in the stem, indicating that indolic and aromatic GSLs have differential tissue profiles compared to aliphatic GSLs. In fact, the aromatic (benzyl and 4-hydroxybenzyl) GSLs are the dominant types of GSLs that make up the largest proportion of the total GSLs in *S. alba*. The level of 4-hydroxybenzyl GSL is much higher than that of benzyl GSL, and both accumulate more in the leaf than in the root (Agerbirk et al., 2008). However, the indolic GSLs (1-methoxy-indol-3ylmethyl, 4-methoxy-indol-3ylmethyl, indol-3ylmethyl, and 4-hydroxyindol-3ylmethyl) are mainly synthesized in hairy root (Kastell et al., 2013). Thus, other unknown enzymes or regulators must exist to control the two pathways separately, aside from the predicted enzymes shared by indolic and aromatic pathways, or some multi-copy genes of these predicted shared enzymes have functionally diverged but cannot be distinguished by only sequence messages.

The GSL core structures subsequently entered side-chain modification processes. The aliphatic methylthioalkyl GSLs were first oxidized to methylsulfinylalkyl GSLs by flavin-monooxygenase glucosinolate S-oxygenase (FMO-GSOX) and then conferred to alkenyl GSLs by alkenyl hydroxalkyl producing (AOP) protein; then, they were consequently decorated with a hydroxyl group by Fe (II)-dependent oxygenase super family protein (GS-OH) to form hydroxylalkenyl GSLs, which is mainly 2-hydroxyl-3buteny GSL in *S. alba* (**Figure 8B**). Four *FMO-GSOXs* (*FMO-GSOX1,2,4,*&*5*), one *AOP* (*AOP1*) and one *GS-OH* were expressed in *S. alba*, and except for *FMO-GSOX4*, all of these genes were significantly expressed more in the root than the stem and the leaf (**Figure 8B,** Table S7). For the indolic pathway, the indolymethyl GSL was catalyzed by CYP81F1 to generate 4-hydroxy-3-indolymethyl GSL. Four transcripts of *CYP81F1* were identified in the transcriptome and were expressed higher in the root (**Figure 8B,** Table S7). For the aromatic pathway, the enzyme catalyzing the benzyl GSL to 4-hydroxybenzyl GSL is still unknown, with GS-OH and CYP81F as possible candidates.

GSLs are stored in vacuoles and will be released and quickly degraded to form isothiocyanates and nitriles when the cells are damaged, such as during food preparation or from pest-chewing. This process is catalyzed by the endogenous plant enzyme myrosinase and can be affected by the epithiospecifier protein (ESP) and reaction environments, such as pH and temperature (Fenwick et al., 1983; Ludikhuyze et al., 2000; Burow et al., 2006; Williams et al., 2008). The endogenous myrosinase hydrolyzes GSLs to isothiocyanates, which gives functions such as anti-insects and anti-microbes in plants and serve as potential anti-tumor compounds in the human diet (Bednarek et al., 2009; Clay et al., 2009; Øverby et al., 2015; Veeranki et al., 2015). When ESP was present, the production of isothiocyanates was reduced, and the GSLs were hydrolyzed to form thiocyanates, epithionitriles, or simple nitriles, depending on the GSL structure (Lambrix et al., 2001; Burow et al., 2009). The biological function of nitriles is still unclear. However, their toxicity effects in human diets and animal feeds are confirmed. Thus, it is important to enhance the isothiocyanate content and reduce the nitrile content in cruciferous crops. In *S. alba*, seven myrosinase and three ESP encoding transcripts were identified (Table S7). Overall, the two gene families were expressed in all of the three organs with no significant tissue bias (**Figure 8B**). However, four myrosinase and

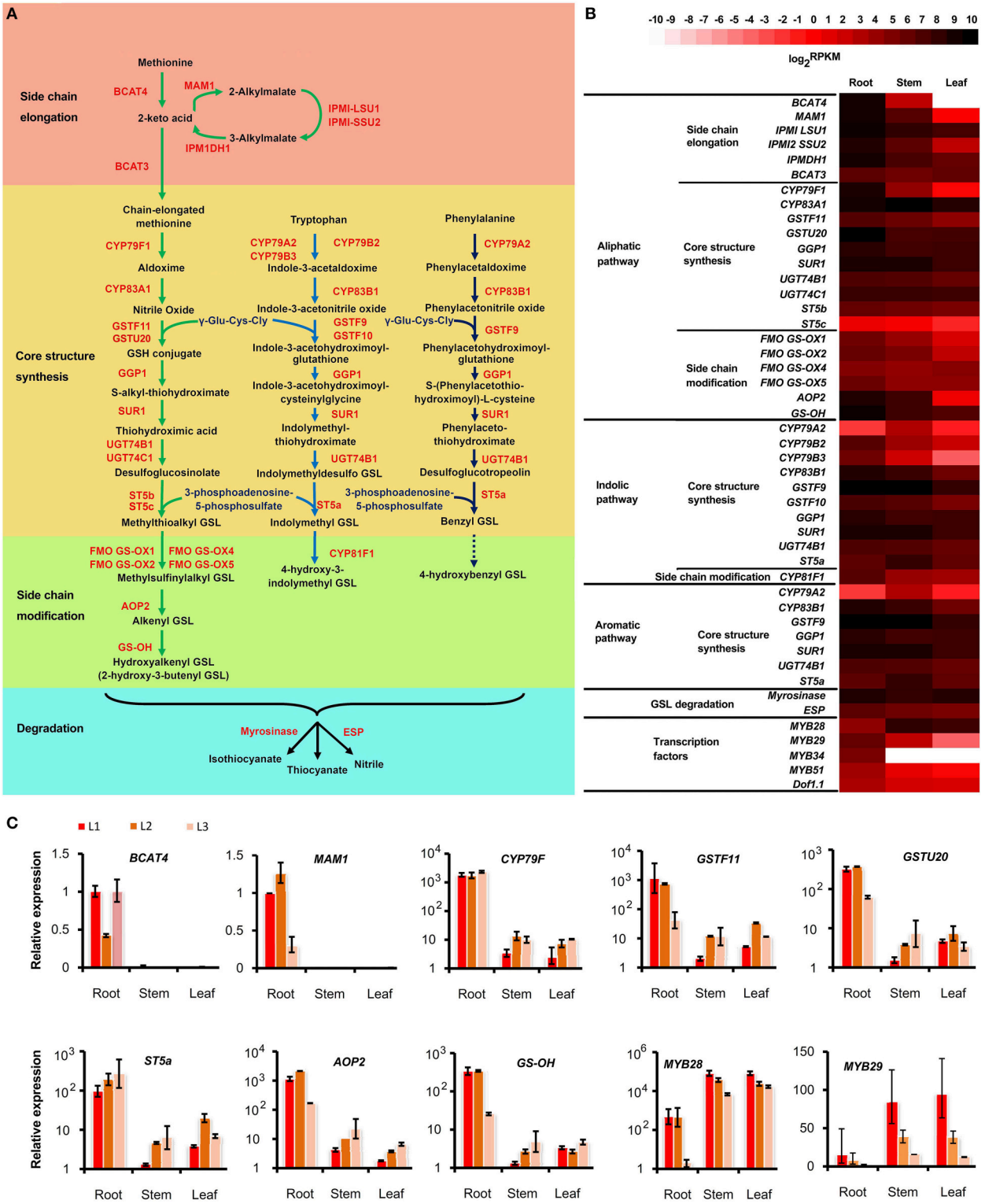


FIGURE 8 | Glucosinolate (GSL) metabolic pathway and its tissue pattern in *Sinapis alba*. **(A)** Pathway map. **(B)** Heatmap showing the expression level of genes encoding the enzymes and regulators in the GSL pathway. **(C)** qPCR results. Bars indicating SDs of technical replicates. BCAT, branched-chain amino acid aminotransferase; MAM, methylthioalkylmalate synthase; IPMI LSU1, Isopropylmalate Isomerase Large Subunit 1; IPMI SSU2, Isopropylmalate Isomerase Small Subunit 2; IPMDH, Isopropylmalate Dehydrogenase; CYP, Cytochromes P450; GSTF, Glutathione S-transferases; GGP, Gamma-Glutamyl Peptidase; SUR, C-S lyase SUPERROOT; UGT, UDP-glucosyltransferase; ST, Desulfoglucosinolate Sulfotransferases; FMO-GSOX, Flavin-Monooxygenase Glucosinolate S-Oxygenase; AOP, Alkenyl hydroxalkyl producing; GS-OH, Fe (II)-dependent oxygenase superfamily protein; ESP, epithiospecifier protein.

one ESP transcripts were specifically expressed in the root, and two myrosinase and one ESP transcripts were predominantly expressed in the stem and the leaf, indicating these genes have been functionally specialized (Table S7). Due to many high GSL-content organs, such as seeds and flowers, not being analyzed, other genes that are undetected in this study could be specifically expressed in these tissues. The multi-copy and tissue specialization properties of these two enzymes offer the possibility to fine-tune the types and amounts of GSL hydrolysis products in the target tissues and desired developmental stages.

The Dof1.1 and six members of the MYB family (MYB28, 29, 34, 51, 76, and 122) transcription factors were reported to regulate the biosynthesis of GSLs (Skirycz et al., 2006; Gigolashvili et al., 2007, 2008; Hirai et al., 2007; Sønderby et al., 2010). From our *S. alba* transcriptome, three *Dof1.1*, one *MYB28*, three *MYB29*, one *MYB34*, and one *MYB51* homologs were found (Table S7). *MYB28* was expressed higher in the stem and the leaf than in the root. *MYB29* and *MYB51* were significantly up-regulated in the root, and *MYB34* was expressed only in the root (**Figure 8B**, Table S7). Research results in *Arabidopsis* have shown that *MYB28* and *MYB29* are key regulators of the aliphatic GSL biosynthesis (Gigolashvili et al., 2007; Hirai et al., 2007; Sønderby et al., 2010). The expression of *MYB29* was well correlated with the tissue pattern of the aliphatic GSL synthesis genes. However, the significant up-expression in the stem and the leaf indicated that MYB28 may have acquired new roles, such as regulating indolic and aromatic GSL biosynthesis in *S. alba*.

To confirm the reliability of the transcriptome profiling, a series of qPCR analyses were performed to 10 genes which showed differential expression between tissues. As shown in **Figure 8C**, 8 of the 10 genes displayed similar expression patters between RNA-Seq and qPCR technology, indicating the relative high quality of the transcriptome profiling. Especially, the extreme low abundant of *BACT4* and *MAM1* in stem and leaf revealed by qPCR analysis of three independent lines strengthened the estimation that the absence of these two genes turned off the biosynthesis of aliphatic GSLs in these tissues. The two conflicting genes, *ST5a* and *MYB29*, might be developmental and environmental sensitive, due to the plants used in the two experiments were grown in two different times and places. The *MYB29* is extreme unstable because its expressions were significantly differed among the three biological replicates.

The Glutathione Metabolic Pathway and Phytochelatin Synthesis

Phytochelatin is a non-mRNA translated glutamylcysteine-repeated peptide (Grill et al., 1989). It is widely believed, that phytochelatin plays an important role in Cd tolerance in plants (Cobbett, 2000). Phytochelatin is synthesized by phytochelatin synthase, with glutathione as building blocks (Grill et al., 1989). Thus, transcripts encoding the phytochelatin synthase and in the glutathione metabolic pathway were identified in the *S. alba* transcriptome (**Figure 9A**). Forty-three transcripts were identified as candidate genes encoding 10 enzymes catalyzing the glutathione metabolic pathway, and two transcripts were isolated as *phytochelatin synthase* genes (Table S8). There is a synthesis-degradation cycle and an oxidation-reduction cycle in this plant (**Figure 9A**). The expression of *gamma-glutamyl transferase* (*GGT*) and *leucyl aminopeptidase* (*pepA*), which were responsible for the degradation of glutathione to L-glutamate, L-cysteine, and glycine, were higher in the root than in the stem and the leaf (**Figures 9A,B**). A similar pattern was also shown for the expression of γ-*glutamylcysteine synthetase* (*gshA*), which catalyzed L-glutamate and L-cysteine to form L-γ-glutamylcysteine (**Figure 9C**). However, the final glutathione biosynthesis step which was catalyzed by glutathione synthetase (gshB) with L-γ-glutamylcysteine and glycine as substrates, showed no significant differences among the three tissues in terms of expressional levels (**Figure 9D**). These results indicated that the glutathione synthesis-degradation cycle was elevated in the root, followed by in the stem, and it was the lowest in the leaf tissues. Degradation products, such as L-glutamate, L-cysteine and glycine, and the metabolic intermediate products, i.e., L-γ-glutamylcysteine, may be shunted to some other metabolic pathways in the root due to the upstream genes being expressed relatively higher in the root but the final step was not significantly different among the three tissues. The transcripts of *phytochelatin synthase* were three times more abundant in the root than the stem and the leaf (**Figure 9E**), indicating that phytochelatin was predominantly synthesized in the root. This phenomenon is in accordance with the assumption that phytochelatin played a key role in Cd detoxification, and its high accumulation in the root resulted in a high level Cd-tolerance in *S. alba*. For the oxidation-reduction cycle, the situation was completely different (**Figure 9A**). The genes encoding glutathione peroxidases (GPXs) were highly expressed in the leaf (**Figure 9F**). However, glutathione reductase (GSR) was not differentially expressed among the three organs (**Figure 9H**). These results indicated that in leaf tissues, glutathione was actively transformed to its oxidized form (GSSG). This transformation could occur because the leaf contains more reductive substances produced by the photosynthesis system. NADPH is a coenzyme for the reduction of GSSG to GSH. The enzymes, isocitrate dehydrogenase (IDH), 6-phosphogluconate dehydrogenase (PGD), and glucose-6-phosphate 1-dehydrogenase (G6PD), catalyzing the reduction of NADP+ to NADPH were expressed higher in the root than in the stem and were expressed the lowest in the leaf (**Figure 9G**), which could drive the reduction process with the corresponding speed in these organs. Finally, glutathione was extensively oxidized in leaf tissues, while reduced forms were more present in the root tissues, and were actively converted into phytochelatin. The expression patterns of the genes in this pathway were replicated and confirmed by qPCR analysis (**Figure 9I**).

CONCLUSIONS

In the present study, a *S. alba* transcriptome was assembled *de novo* for the first time, to the best of our knowledge. The 47,972 generated unigenes were a mean length of 1185 nt and

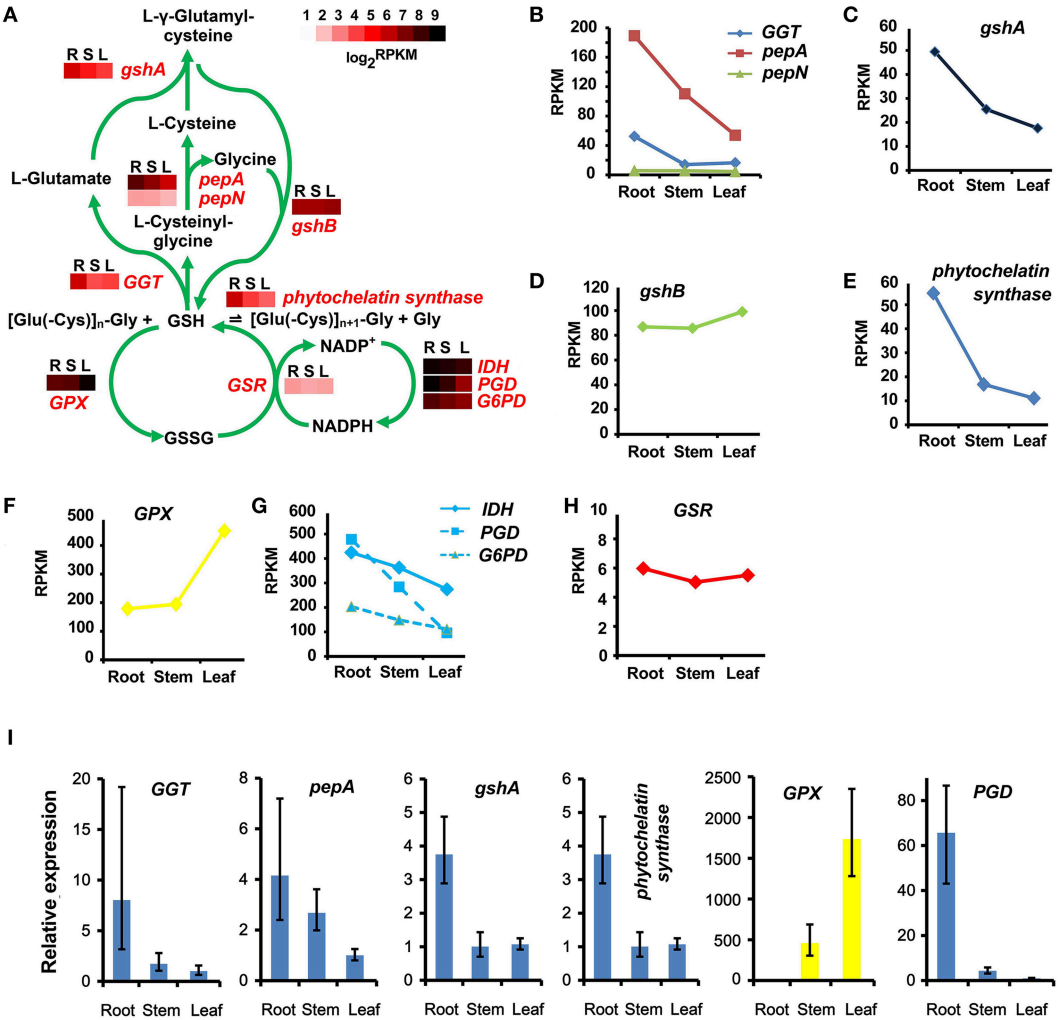


FIGURE 9 | Glutathione and phytochelatin synthetic pathways and their tissue patterns in *Sinapis alba*. (A) The glutathione cycling and phytochelatin synthetic pathway map with heatmaps showing the expression levels of the corresponding enzymes. R, root; S, stem; L, leaf. **(B–H)** Line charts showing the gene expression levels based on RNA-seq. **(I)** qPCR results. Bars indicating SDs of three biological and two technical replicates. GGT, gamma-glutamyl transferase; pepA, leucyl aminopeptidase; pepN, aminopeptidase; gshA, γ-glutamylcysteine synthetase; gshB, glutathione synthetase; GPX, glutathione peroxidase; IDH, isocitrate dehydrogenase [NADP]; PGD, 6-phosphogluconate dehydrogenase; G6PD, glucose-6-phosphate 1-dehydrogenase; GSR, glutathione reductase.

had an N50 of 1672 nt, indicating a high quality reference transcriptome for genetic studies. The produced 14,727 SSRs will be useful for the genetic analysis of this non-model crop. Although no reference genome was available, 97% of the unigenes were functionally annotated. Expression profiles showed that the root accumulated the largest fraction of specifically and predominantly expressed genes, indicating its involvement in many more specialized functions. The genes predominantly expressed in the root were enriched in pathways related to lignin and cellulose syntheses, plant-pathogen interactions, and pathways potentially responsible for heavy metal chelating and detoxification. The glucosinolate and phytochelatin metabolic pathways, which confer the characteristics and utilities of this plant, were intensively analyzed. The genes encoding aliphatic GSLs were predominantly expressed in the root. The absence of aliphatic GSLs in leaf tissues was most likely due to the lack of *BCAT4* expression and the low expressions of *MAM1* and *CYP79F1*, which efficiently blocked the pathway. Glutathione in the root was extensively converted into phytochelatin, but in the leaf, it was actively converted to its oxidized form. The transcriptome and SSR markers from this study will benefit basic research on and the molecular breeding of *S. alba* and will also be useful for studying the mechanisms of GSLs, phytoremediation and other important traits, as well as the transfer of these beneficial traits to other crops.

AUTHOR CONTRIBUTIONS

XZ and XL designed the study. XZ, TL, MD, and JS performed the experiments. XZ analyzed the data and drafted the manuscript. All of the authors carefully checked and approved this manuscript.

DATA ACCESS

RNAseq data are available at EMBL/NCBI/SRA (accession numbers SRR2961888, SRR2961889, and SRR2961890).

ACKNOWLEDGMENTS

This work was supported by grants from the National Key Technology R&D Program of the Ministry of Science and Technology of China (2013BAD01B04-2), and the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fpls.2016.00259

REFERENCES

Abdi, H. (2007). "Bonferroni and Sidák corrections for multiple comparisons," in *Encyclopedia of Measurement and Statistics*, ed *N. J. Salkind (Thousand Oaks, CA: Sage)*, 103–107.

Abdull Razis, A. F., De Nicola, G. R., Pagnotta, E., Iori, R., and Ioannides, C. (2012). 4-Methylsulfanyl-3-butenyl isothiocyanate derived from glucoraphasatin is a potent inducer of rat hepatic phase II enzymes and a potential chemopreventive agent. *Arch. Toxicol.* 86, 183–194. doi: 10.1007/s00204-011-0750-x

Agerbirk, N., Warwick, S. I., Hansen, P. R., and Olsen, C. E. (2008). *Sinapis* phylogeny and evolution of glucosinolates and specific nitrile degrading enzymes. *Phytochemistry* 69, 2937–2949. doi: 10.1016/j.phytochem.2008.08.014

Bednarek, P., Pislewska-Bednarek, M., Svatos, A., Schneider, B., Doubsky, J., Mansurova, M., et al. (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* 323, 101–106. doi: 10.1126/science.1163732

Bodnaryk, R. P., and Lamb, R. J. (1991). Mechanisms of resistance to the flea beetle, *Phyllotreta cruciferae* (Goeze), in mustard seedlings, *Sinapis alba* L. *Can. J. Plant Sci.* 71, 13–20. doi: 10.4141/cjps91-002

Brown, J., Brown, A. P., Davis, J. B., and Erickson, D. (1997). Intergeneric hybridization between *Sinapis alba* and *Brassica napus*. *Euphytica* 93, 163–168. doi: 10.1023/A:1002905816887

Burow, M., Losansky, A., Müller, R., Plock, A., Kliebenstein, D. J., and Wittstock, U. (2009). The genetic basis of constitutive and herbivore-induced ESP-independent nitrile formation in *Arabidopsis*. *Plant Physiol.* 149, 561–574. doi: 10.1104/pp.108.130732

Burow, M., Markert, J., Gershenzon, J., and Wittstock, U. (2006). Comparative biochemical characterization of nitrile-forming proteins from plants and insects that alter myrosinase-catalysed hydrolysis of glucosinolates. *FEBS J.* 273, 2432–2446. doi: 10.1111/j.1742-4658.2006.05252.x

Clay, N. K., Adio, A. M., Denoux, C., Jander, G., and Ausubel, F. M. (2009). Glucosinolate metabolites required for an *Arabidopsis* innate immune response. *Science* 323, 95–101. doi: 10.1126/science.1164627

Cobbett, C. S. (2000). Phytochelatins and their roles in heavy metal detoxification. *Plant Physiol.* 123, 825–832. doi: 10.1104/pp.123.3.825

Dong, C. H., Li, C., Yan, X. H., Huang, S. M., Huang, J. Y., Wang, L. J., et al. (2012). Gene expression profiling of *Sinapis alba* leaves under drought stress and rewatering growth conditions with Illumina deep sequencing. *Mol. Biol. Rep.* 39, 5851–5857. doi: 10.1007/s11033-011-1395-9

Fahey, J. W., Zalcmann, A. T., and Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51. doi: 10.1016/S0031-9422(00)00316-2

Fenwick, G. R., Heaney, R. K., and Mullin, W. J. (1983). Glucosinolates and their breakdown products in food and food plants. *Crit. Rev. Food Sci.* 18, 123–201. doi: 10.1080/10408398209527361

Gigolashvili, T., Engqvist, M., Yatusevich, R., Müller, C., and Flugge, U. I. (2008). HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. *New Phytol.* 177, 627–642. doi: 10.1111/j.1469-8137.2007.02295.x

Gigolashvili, T., Yatusevich, R., Berger, B., Müller, C., and Flugge, U. I. (2007). The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant J.* 51, 247–261. doi: 10.1111/j.1365-313X.2007.03133.x

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Grill, E., Löffler, S., Winnacker, E. L., and Zenk, M. H. (1989). Phytochelatins, the heavy-metal-binding peptides of plants, are synthesized from glutathione by a specific gamma-glutamylcysteine dipeptidyl transpeptidase (*phytochelatin synthase*). *Proc. Natl. Acad. Sci. U.S.A.* 86, 6838–6842. doi: 10.1073/pnas.86.18.6838

Grubb, C. D., and Abel, S. (2006). Glucosinolate metabolism and its control. *Trends Plant Sci.* 11, 89–100. doi: 10.1016/j.tplants.2005.12.006

Hall, J. L. (2002). Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.* 53, 1–11. doi: 10.1093/jexbot/53.366.1

Hemingway, J. S. (1995). "The mustard species: condiment and food ingredient use and potential as oilseed crops," in *Brassica Oilseeds Production and Utilization*, eds D. *Kimber and D. I. McGregor (Wallingford, CT: CAB International)*, 373–383.

Hirai, M. Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., et al. (2007). Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 6478–6483. doi: 10.1073/pnas.0611629104

Iseli, C., Jongeneel, C. V., and Bucher, P. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 138–148.

Javidfar, F., and Cheng, B. (2013). Construction of a genetic linkage map and QTL analysis of erucic acid content and glucosinolate components in yellow mustard (*Sinapis alba* L.). *BMC Plant Biol.* 13:142. doi: 10.1186/1471-2229-13-142

Jiang, J., Shao, Y., Li, A., Zhang, Y., Wei, C., and Wang, Y. (2013). FT-IR and NMR study of seed coat dissected from different colored progenies of *Brassica napus-Sinapis alba* hybrids. *J. Sci. Food Agri.* 93, 1898–1902. doi: 10.1002/jsfa.5986

Kastell, A., Smetanska, I., Ulrichs, C., Cai, Z., and Mewis, I. (2013). Effects of phytohormones and jasmonic acid on glucosinolate content in hairy root cultures of *Sinapis alba* and *Brassica rapa*. *Appl. Biochem. Biotech.* 169, 624–635. doi: 10.1007/s12010-012-0017-x

Lambrix, V., Reichelt, M., Mitchell-Olds, T., Kliebenstein, D. J., and Gershenzon, J. (2001). The *Arabidopsis* epithiospecifier protein promotes the hydrolysis of glucosinolates to nitriles and influences *Trichoplusia ni* herbivory. *Plant Cell* 13, 2793–2807. doi: 10.1105/tpc.13.12.2793

Lee, R. W., Malchev, I. T., Rajcan, I., and Kott, L. S. (2014). Identification of putative quantitative trait loci associated with a flavonoid related to resistance to cabbage seedpod weevil (*Ceutorhynchus obstrictus*) in canola derived from an intergeneric cross, *Sinapis alba* x *Brassica napus*. *Theor. Appl. Genet.* 127, 419–428. doi: 10.1007/s00122-013-2228-0

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A., et al. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nat. Commun.* 5, 3930. doi: 10.1038/ncomms4930

Ludikhuyze, L., Rodrigo, L., and Hendrickx, M. (2000). The activity of myrosinase from broccoli (*Brassica oleracea* L. cv. Italica): influence of intrinsic and extrinsic factors. *J. Food Protect.* 63, 400–403.

Mendoza-Cázatl, D. G., Jobe, T. O., Hauser, F., and Schroeder, J. I. (2011). Long-distance transport, vacuolar sequestration, tolerance, and transcriptional responses induced by cadmium and arsenic. *Curr. Opin. Plant Biol.* 14, 554–562. doi: 10.1016/j.pbi.2011.07.004

Mitsui, Y., Shimomura, M., Komatsu, K., Namiki, N., Shibata-Hatta, M., Imai, M., et al. (2015). The radish genome and comprehensive gene expression

profile of tuberous root formation and development. *Sci. Rep.* 5:10835. doi: 10.1038/srep10835

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5, 621–628. doi: 10.1038/nmeth.1226

Øverby, A., Bævre, M. S., Thangstad, O. P., and Bones, A. M. (2015). Disintegration of microtubules in *Arabidopsis thaliana* and bladder cancer cells by isothiocyanates. *Front. Plant Sci.* 6:6. doi: 10.3389/fpls.2015.00006

Peng, C., Zhao, S. Q., Zhang, J., Huang, G. Y., Chen, L. Y., and Zhao, F. Y. (2014). Chemical composition, antimicrobial property and microencapsulation of Mustard (*Sinapis alba*) seed essential oil by complex coacervation. *Food Chem.* 165, 560–568. doi: 10.1016/j.foodchem.2014.05.126

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., et al. (2003). TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19, 651–652. doi: 10.1093/bioinformatics/btg034

Pino Del Carpio, D., Basnet, R. K., Arends, D., Lin, K., De Vos, R. C., Muth, D., et al. (2014). Regulatory network of secondary metabolism in *Brassica rapa*: insight into the glucosinolate pathway. *PLoS ONE* 9:e107123. doi: 10.1371/journal.pone.0107123

Plociniczak, T., Kukla, M., Watroba, R., and Piotrowska-Seget, Z. (2013). The effect of soil bioaugmentation with strains of Pseudomonas on Cd, Zn and Cu uptake by *Sinapis alba* L. *Chemosphere* 91, 1332–1337. doi: 10.1016/j.chemosphere.2013.03.008

Popova, I. E., and Morra, M. J. (2014). Simultaneous quantification of sinigrin, sinalbin, and anionic glucosinolate hydrolysis products in *Brassica juncea* and *Sinapis alba* seed extracts using ion chromatography. *J. Agr. Food Chem.* 62, 10687–10693. doi: 10.1021/jf503755m

Romualdi, C., Bortoluzzi, S., D'Alessi, F., and Danieli, G. A. (2003). IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genomics* 12, 159–162. doi: 10.1152/physiolgenomics.00096.2002

Skirycz, A., Reichelt, M., Burow, M., Birkemeyer, C., Rolcik, J., Kopka, J., et al. (2006). DOF transcription factor AtDof1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in *Arabidopsis*. *Plant J.* 47, 10–24. doi: 10.1111/j.1365-313X.2006.02767.x

Sønderby, I. E., Burow, M., Rowe, H. C., Kliebenstein, D. J., and Halkier, B. A. (2010). A complex interplay of three R2R3 MYB transcription factors determines the profile of aliphatic glucosinolates in *Arabidopsis*. *Plant Physiol.* 153, 348–363. doi: 10.1104/pp.109.149286

Thompson, K. F. (1963). Resistance to the cabbage aphid (*Brevicoryne brassicae*) in *Brassica* plants. *Nature* 198, 209. doi: 10.1038/198209a0

Tian, E., Zeng, F., MacKay, K., Roslinsky, V., and Cheng, B. (2014). Detection and molecular characterization of two *FAD3* genes controlling linolenic acid content and development of allele-specific markers in yellow mustard (*Sinapis alba*). *PLoS ONE* 9:e97430. doi: 10.1371/journal.pone.0097430

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3–new capabilities and interfaces. *Nucleic Acids Res.* 40, e115. doi: 10.1093/nar/gks596

Vastenhout, K. J., Tornberg, R. H., Johnson, A. L., Amolins, M. W., and Mays, J. R. (2014). High-performance liquid chromatography-based method to evaluate kinetics of glucosinolate hydrolysis by *Sinapis alba* myrosinase. *Anal. Biochem.* 465, 105–113. doi: 10.1016/j.ab.2014.07.017

Veeranki, O. L., Bhattacharya, A., Tang, L., Marshall, J. R., and Zhang, Y. (2015). Cruciferous vegetables, isothiocyanates, and prevention of bladder cancer. *Curr. Pharmacol. Rep.* 1, 272–282. doi: 10.1007/s40495-015-0024-z

Wang, H., Wu, J., Sun, S., Liu, B., Cheng, F., Sun, R., et al. (2011). Glucosinolate biosynthetic genes in *Brassica rapa*. *Gene* 487, 135–142. doi: 10.1016/j.gene.2011.07.021

Wiesner, M., Schreiner, M., and Zrenner, R. (2014). Functional identification of genes responsible for the biosynthesis of 1-methoxy-indol-3-ylmethyl-glucosinolate in *Brassica rapa* ssp. *chinensis*. *BMC Plant Biol.* 14:124. doi: 10.1186/1471-2229-14-124

Williams, D. J., Critchley, C., Pun, S., Nottingham, S., and O'Hare, T. J. (2008). Epithiospecifier protein activity in broccoli: the link between terminal alkenyl glucosinolates and sulphoraphane nitrile. *Phytochemistry* 69, 2765–2773. doi: 10.1016/j.phytochem.2008.09.018

Wittstock, U., and Halkier, B. A. (2002). Glucosinolate research in the *Arabidopsis* era. *Trends Plant Sci.* 7, 263–270. doi: 10.1016/S1360-1385(02)02273-2

Zang, Y. X., Kim, H. U., Kim, J. A., Lim, M. H., Jin, M., Lee, S. C., et al. (2009). Genome-wide identification of glucosinolate synthesis genes in *Brassica rapa*. *FEBS J.* 276, 3559–3574. doi: 10.1111/j.1742-4658.2009.07076.x

Zielniok, K., Szkoda, K., Gajewska, M., and Wilczak, J. (2015). Effect of biologically active substances present in water extracts of white mustard and coriander on antioxidant status and lipid peroxidation of mouse C2C12 skeletal muscle cells. *J. Anim. Physiol. Anim. Nutr.* doi: 10.1111/jpn.12412

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Zhang, Liu, Duan, Song and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership