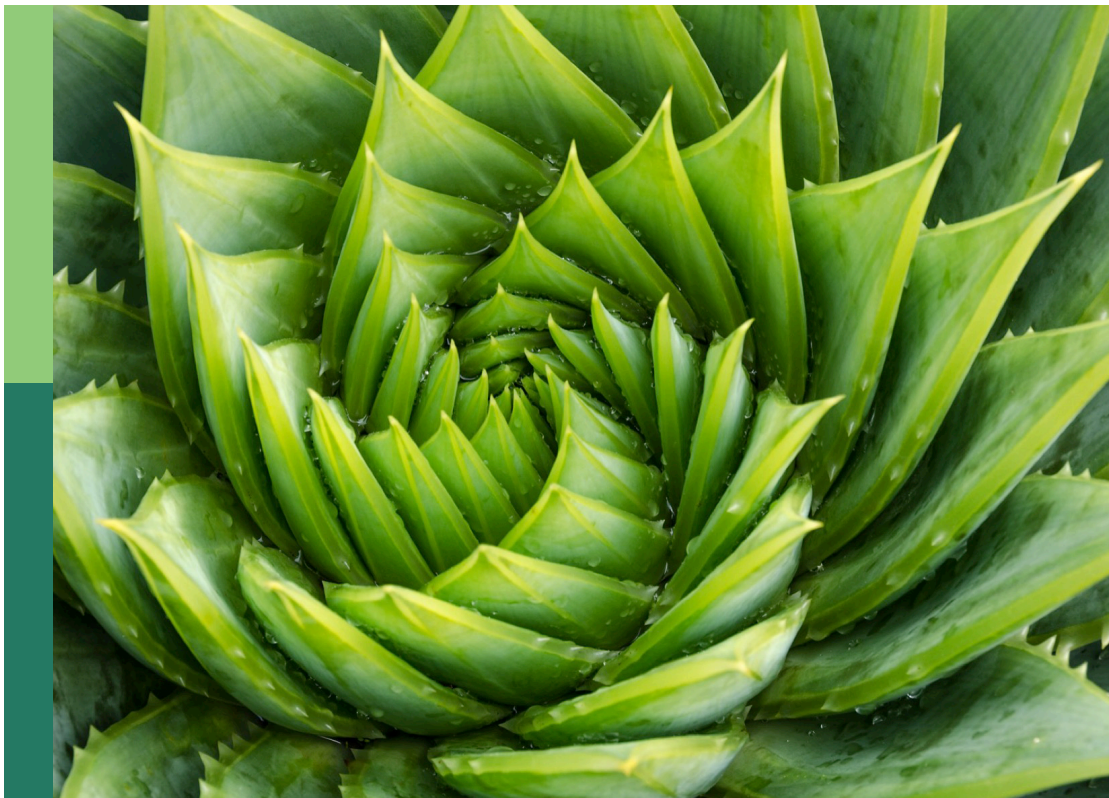
frontiers | Research Topics

Recent advances in flower and fruit development in perennial plants

Edited by
Shunquan Lin, Hao Yu, Zhongchi Liu and Yuexue Liu

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-2989-8
DOI 10.3389/978-2-8325-2989-8

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: frontiersin.org/about/contact

Recent advances in flower and fruit development in perennial plants

Topic editors

Shunquan Lin — South China Agricultural University, China
Hao Yu — National University of Singapore, Singapore
Zhongchi Liu — University of Maryland, College Park, United States
Yuexue Liu — Shenyang Agricultural University, China

Citation

Lin, S., Yu, H., Liu, Z., Liu, Y., eds. (2023). *Recent advances in flower and fruit development in perennial plants*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2989-8

Table of contents

frontiers in Plant Science

ORIGINAL RESEARCH
published: 31 December 2021
doi: 10.3389/fpls.2021.816086

EjRAV1/2 Delay Flowering Through Transcriptional Repression of *EjFTs* and *EjSOC1s* in Loquat

Ze Peng[1†], Man Wang[1†], Ling Zhang[1,5], Yuanyuan Jiang[1], Chongbin Zhao[1], Muhammad Qasim Shahid[1,4], Yunlu Bai[1], Jingjing Hao[1], Jiangrong Peng[1], Yongshun Gao[3], Wenbing Su[1,2,*] and Xianghui Yang[1]**

[1] State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, Key Laboratory of Innovation and Utilization of Horticultural Crop Resources in South China (Ministry of Agriculture and Rural Affairs), South China Agricultural University, Guangzhou, China, [2] Fruit Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China, [3] Beijing Academy of Forestry and Pomology Sciences, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [4] College of Agriculture, South China Agricultural University, Guangzhou, China, [5] Lushan Botanical Garden Jiangxi Province and Chinese Academy of Sciences, Lushan, China

OPEN ACCESS

Edited by:
Zhongchi Liu,
University of Maryland, College Park, United States

Reviewed by:
Xueren Yin,
Zhejiang University, China
Tianchi Wang,
The New Zealand Institute for Plant and Food Research Ltd., New Zealand

***Correspondence:**
Xianghui Yang
gzyxh@scau.edu.cn
Wenbing Su
suwenbing13@163.com
Yongshun Gao
yongshungao@163.com

[†] These authors have contributed equally to this work and share first authorship

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

Received: 16 November 2021
Accepted: 06 December 2021
Published: 31 December 2021

Citation:
Peng Z, Wang M, Zhang L, Jiang Y, Zhao C, Shahid MQ, Bai Y, Hao J, Peng J, Gao Y, Su W and Yang X (2021) EjRAV1/2 Delay Flowering Through Transcriptional Repression of EjFTs and EjSOC1s in Loquat. Front. Plant Sci. 12:816086. doi: 10.3389/fpls.2021.816086

Most species in Rosaceae usually need to undergo several years of juvenile phase before the initiation of flowering. After 4–6 years' juvenile phase, cultivated loquat (*Eriobotrya japonica*), a species in Rosaceae, enters the reproductive phase, blooms in the autumn and sets fruits during the winter. However, the mechanisms of the transition from a seedling to an adult tree remain obscure in loquat. The regulation networks controlling seasonal flowering are also largely unknown. Here, we report two *RELATED TO ABI3 AND VP1* (*RAV*) homologs controlling juvenility and seasonal flowering in loquat. The expressions of *EjRAV1/2* were relatively high during the juvenile or vegetative phase and low at the adult or reproductive phase. Overexpression of the two *EjRAV*s in *Arabidopsis* prolonged (about threefold) the juvenile period by repressing the expressions of flowering activator genes. Additionally, the transformed plants produced more lateral branches than the wild type plants. Molecular assays revealed that the nucleus localized EjRAVs could bind to the CAACA motif of the promoters of flower signal integrators, *EjFT1/2*, to repress their expression levels. These findings suggest that *EjRAVs* play critical roles in maintaining juvenility and repressing flower initiation in the early life cycle of loquat as well as in regulating seasonal flowering. Results from this study not only shed light on the control and maintenance of the juvenile phase, but also provided potential targets for manipulation of flowering time and accelerated breeding in loquat.

Keywords: flowering time, juvenility, annual flowering, FT, branching, loquat, RAV

INTRODUCTION

Flowering is a crucial process required for reproductive success, integrating external and internal signals (Song et al., 2013). This process is energy-consuming and therefore plants need to undergo diverse juvenile periods to accumulate sufficient reserves before flowering (Andrés and Coupland, 2012). For annual plants, such as *Arabidopsis*, the whole life cycle is limited to several months

with a short juvenile phase followed by flowering and fruit setting in the adult phase, and the life cycle ends as fruits become mature (Somerville and Koornneef, 2002). In contrast, the long-lived perennial trees, such as poplar and eucalyptus, have a long juvenile phase that lasts for several years (Missiaggia et al., 2005; Hsu et al., 2006). Although perennial plants bloom annually or seasonally after the initiation of the first flowering, the long juvenile period is time-consuming and needs to be shortened to accelerate breeding and other research programs for economically important perennial plants (Watson et al., 2018).

As a key developmental event of a plant's life cycle, the juvenile-to-adult transition of both annual *Arabidopsis* and perennial trees is regulated by gradually decreasing the *miR156/157* expressions as well as increasing the expressions of its target genes encoding a set of plant specific SPL (SQUAMOSA PROMOTER BINDING PROTEIN-LIKE) proteins (Xu et al., 2016; Jia et al., 2017). Commonly, the miR156 levels would decline as the plant age increases, which elevates the expression levels of SPL transcription factors to promote flowering by inducing the expressions of floral integrator genes including *FLOWERING LOCUS T* (*FT*), *SUPPRESSOR OF OVEREXPRESSION OF CO1* (*SOC1*), *APETALA1* (*AP1*), and *LEAFY* (*LFY*) (Wang, 2014; Xu et al., 2018).

The Rosaceae family possesses approximately 90 genera and 3,000 species with several fruit trees like apple, pear, *Prunus,* and strawberry producing delicious and nutritious fruits; ornamental plants such as rose, *Prunus mume* and oriental cherry as well as medicinal plants like loquat (also a fruit tree), *Potentilla chinensis* and hawthorn (Potter et al., 2007; Xiang et al., 2016). The growth habits vary greatly in this family, from herbaceous rosette plants to deciduous or evergreen trees. For example, the herbaceous strawberry could flower in its first season, while woody fruit trees usually need several years to end the juvenile period (Kurokura et al., 2013). FT, a major component of florigen, and its antagonistic homolog TERMINAL FLOWER 1 (TFL1) are vital regulators of plant flowering (Daniel, 2015), and their homologs in Rosaceae have proven to initiate the reproductive phase and regulate seasonal flowering (Kotoda et al., 2010; Kurokura et al., 2013). Overexpression of FT orthologs derived from fruit tree species in Arabidopsis induced its early flowering (Tränkner et al., 2010; Yarur et al., 2016). Currently available studies in fruit trees suggested the conserved role of FT in promoting the floral transition, and additional functions were implied in these studies based on the observations of pleiotropic transgenic phenotypes (Goeckeritz and Hollender, 2021). In some other fruit species such as kiwifruit, the FT orthologs were reported to play a role in maturity regulation and vegetative phenology in addition to flowering (Voogd et al., 2017). As another key flowering integrator, SOC1 was frequently reported to regulate the endodormancy break (Gómez-Soto et al., 2021). In sweet cherry, the *PavSOC1* gene was reported to interact with *dormancy-associated MADS-box* (*DAM*) genes to co-regulate flower development (Wang J. et al., 2020). In peach (*Prunus persica*), *PpSOC1* has been used as a marker for the timing of endodormancy break (Halász et al., 2021). Meanwhile, efforts have been made to use these genes as targets to shorten the juvenile period and accelerate breeding (Yamagishi et al., 2016; McGarry et al., 2017; Li et al., 2019). Nevertheless, the gene network regulating the initiation of flowering in Rosaceae is still rarely reported, resulting in limited novel gene targets.

Loquat (*Eriobotrya japonica*) is a subtropical evergreen fruit tree native to China (Su et al., 2021). The seedlings need to undergo 4–6 years of juvenile phase before the first flowering (Zheng et al., 1991). Its deciduous relatives, such as apple and pear, induce inflorescence in early summer, develop flower meristems in autumn, enter endodormancy in late autumn, and finally bloom in the next spring (Kurokura et al., 2013). By contrast, the cultivated loquat induces inflorescence and develops flower in the same year (Jiang et al., 2019), but the process from flowering to fruit setting occurs quickly in the winter season and fruits ripen in spring. The life history and annual growth cycle of cultivated loquat were illustrated in **Supplementary Figure 1**. As for loquat, previous works from our group and other researchers have identified some activators including *EjAP1* (Liu et al., 2013), *EdFT* and *EdFD* (Zhang et al., 2016), *EjLFY* (Liu et al., 2017), *EjSOC1* (Jiang et al., 2019), *EdCO* and *EdGI* (Zhang et al., 2019) and *EjSPL3/4/5/9* (Jiang et al., 2020a), as well as repressors like *EjTFL1* (Jiang et al., 2020b) and *EjFRI* (Chen W. et al., 2020). The functions of these genes were conserved between Arabidopsis and loquat. However, whether these genes or other genes in loquat have any relation with the control or maintenance of juvenility is still unknown. Meanwhile, the pathways of these genes to induce flowering remain to be elucidated.

RELATED TO ABI3 AND VP1 (*RAV*) family genes, including *TEMPRANILLO* (*TEM*), play vital roles in linking photoperiod-, gibberellin- as well as other pathways to control flowering and juvenility in *Arabidopsis* (Castillejo and Pelaz, 2008; Osnato et al., 2012; Matias-Hernandez et al., 2014). However, it is unknown whether the homologs of *RAV* modulate flower initiation in long-lived perennial trees. To test this idea, here we cloned two *RAV* homologs (*EjRAV1* and *EjRAV2*) from *E. japonica*. *EjRAV1* and *EjRAV2* turned out to be good biomarkers that maintain juvenility and vegetative stages in loquat with abundant transcripts at these stages and rarely expressed at reproductive stages. In this study, we examined the roles of *EjRAV1* and *EjRAV2* in the transition from juvenile phase to reproductive phase in loquat, and investigated their functions in seasonal flowering in trees during fruit production. Our work revealed that EjRAV proteins could bind to the CAACA motifs in promoter regions of *FT* homologs to repress flower signal integration in loquat.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

According to our previous observations, the cultivated loquat trees start flower primordial differentiation and inflorescence formation from late June to July (Liu et al., 2013; Jiang et al., 2019). Hence, shoot apical meristem (SAM) observation as well as the collection of SAM and mature leaf samples of "Zaozhong-6," a main cultivar in China, were performed from 16th May to 8th August in 2018 at the loquat germplasm resource

preservation garden (South China Agricultural University, Guangzhou, China). The samples were collected at around 10:00–11:00 a.m. in every 7 days (except for 23rd May) during this time window. To decipher the correlations of gene expression levels between *EjRAV1/2* and other flowering regulating genes for the duration of juvenile phase, mature leaves from 1-year-old seedlings, 2-year-old seedlings, 2-year-old grafted trees and fruit-bearing adult trees (all trees with "Zaozhong-6" background) were collected on 11th July, 2018 (at around 10:00–11:00 a.m.) with three biological replicates from our germplasm resource preservation garden. For each replicate, three mature leaves were randomly collected from one plant and were immediately frozen in liquid nitrogen.

Arabidopsis and *Nicotiana benthamiana* were used for stable and transient genetic transformation, respectively. Columbia-0 (Col-0) ecotype *Arabidopsis* and *N. benthamiana* were preserved by our research group, and *tem2* (Col-0 background) mutant was purchased from the *Arabidopsis* Biological Resource Center[1] with a stock number of CS909678. All these plants were planted in a long-day (16 h light/8 h dark) greenhouse at 22°C. The 4-week-old T3 *Arabidopsis* plants were collected for gene expression assays.

RAV Gene Family Identification in Loquat Genome

The six RAV family protein sequences of *Arabidopsis* were downloaded from GenBank and used as queries to Blast search for homologs in "Jiefangzhong" loquat genome (Su et al., 2021) using TBtools software with an *E*-value cutoff at $1e^{-10}$, protein identity≥40 and bit score≥100 (Chen C. et al., 2020). Expression level heat maps of the six identified *EjRAV*s in 10 tissues were drawn based on FPKM values from transcriptome data (available at China National GeneBank database with accession number CNP0001531) using TBtools software (Chen C. et al., 2020). Multiple sequence alignment was performed using ClustalX[2]. The phylogenetic relationship of the *RAV* genes was calculated by MEGA5 (Tamura et al., 2011), with the bootstrap values at the branch nodes calculated from 100 replications. The following proteins were used for phylogenetic analysis: AtTEM/EDF1 (NP_173927.1), AtRAV2/EDF2/TEM2 (NP_173927.1), AtE DF3 (NP_189201.1), AtRAV1/EDF4 (NP_172784.1), AtRA V3 (CAA0284621.1), AtRAVL3 (NP_175524.1); CaRAV (AAQ05799.1), CsRAV1 (AEZ02303.1), DkERF32 (QFU8520 5.1), DkERF34 (QFU85207.1), EjRAV1 (QBQ58100.1), EjRAV2 (QBQ58101.1), FaRAV1 (AZL19498.1), FaRAV3 (AZL19 500.1), FaRAV5 (AZL19502.1), FaRAV6 (AZL19503.1), GmRAV (NP_001237600.1), MdRAV1 (MD13G1046100), MdRAV2 (AU Z96416.1), MtRAV1 (XP_003591822.1), MtRAV2 (AES97413.2), MtRAV3 (AES63108.1), NtRAV (NP_001311676.1), OsRAVL1 (ADJ57333.1), OsRAV6 (XP_015624169.1), PtRAV1 (XP_00 2315958.2), PtRAV2 (XP_002311438.2), PtRAV3 (XP_00230 4025.1), SlRAV2 (ABY57635.1). AtDRN-LIKE (NP_173864.1) was selected as an out-group.

[1]https://abrc.osu.edu

[2]http://www.ebi.ac.uk

RNA Extraction and cDNA Preparation

Total RNA of loquat leaves and *Arabidopsis* plants were extracted with the EASYspin Plus plant RNA extraction kit (Aidlab, China) under the manufacturers' instructions. Then, the PrimeScript™ RT reagent Kit (TaKaRa, Japan) was used to prepare the first-strand cDNA of all the samples.

Gene Transformation in *Arabidopsis*

EjRAV1 and *EjRAV2* sequences were amplified and cloned into the pGreen-35S vector to construct the pGreen-*35S:EjRAV1* and pGreen-*35S:EjRAV2* vectors with the primer pairs listed in **Supplementary Table 1** using the PrimeSTAR® HS DNA Polymerase (TaKaRa, Japan). The PCR amplicons and linearized vector (digested by *Hin*dIII and *Eco*RI; New England Biolabs, United States) were fused with a Hieff Clone® Plus One Step Cloning Kit (Yeasen Biotech, China). *35S:EjRAV1* and *35S:EjRAV2* vectors were first transformed into the *Agrobacterium tumefaciens* GV3101:psoup cells, then floral-dip method was used to transfer overexpression vectors into *Arabidopsis* as previously described (Henriques et al., 2006).

Quantitative Real-Time PCR Assays

Gene expression levels were analyzed using quantitative real-time PCR as previously done (Su et al., 2019). The specific quantitative primers were designed using the Primer Premier 6.0 software[3]. *EjRPL18* (MH196507) and *AtUBQ10* (AL161503) were used as the reference genes for loquat and *Arabidopsis* gene expression assays. Detailed primer sequence information for all the detected genes were listed in **Supplementary Table 1**. All the biological samples were detected in triplicates using iTaq™ universal SYBR Green Supermix (Bio-Rad, United States) in the LightCycler480 system (Roche, Sweden).

Transient Expression Assays in *Nicotiana benthamiana*

The coding regions without stop codon of *EjRAV1* and *EjRAV2* were fused into pGreen-35S-GFP (green fluorescent protein) with primer pairs listed in **Supplementary Table 1**. The vectors were then transformed into *Agrobacterium tumefaciens* GV3101:psoup, GV3101:psoup cells at OD_{600} = 1.0 were centrifuged at 4,500 g for 5 min, and the pellets were suspended with MMA solution (MES, $MgCl_2$ and Acetosyringone) to OD_{600} = 0.2 (Hellens et al., 2005). The dilute *Agrobacterium* cells were then incubated in dark under room temperature for 1 h and finally used to inject into *N. benthamiana* leaves. The tobacco leaves were incubated in long day greenhouse for 2 days, images of EjRAV1-GFP, EjRAV1-GFP and 35S-GFP were captured with the Observer. D1 fluorescence microscope system (Carl Zeiss, Germany).

The pGreen-0800:*EjSOC1* vectors were previously prepared (Jiang et al., 2020a). *EjFT1* and *EjFT2* promoter sequences were fused into pGreen-0800 vector with primer pairs listed in **Supplementary Table 1**, and the protocol for pGreen-0800

[3]http://www.premierbiosoft.com/primerdesign/

reporter vector construction was similar to that used in pGreen-*35S:EjRAV1* vector preparation. Then, PLACE website[4] was used to check the binding sites for RAV proteins on each promoter. The positions of predicted RAV binding sites on these promoters were highlighted in **Supplementary Figure 2** and listed in the **Supplementary Table 2**.

The coding regions of *EjRAV1* and *EjRAV2* were fused into the pSAK277 vector to construct the effector vector (Hellens et al., 2005) with primer pairs listed in **Supplementary Table 1**. The effector and reporter vectors were then transformed into *Agrobacterium tumefaciens* GV3101:psoup cells. Then, the mixed *Agrobacterium* solution with effector and reporter vectors was injected into *N. benthamiana* leaves as formerly performed (Hellens et al., 2005). The activities of firefly and renilla dual-luciferase were detected using a Dual-Luciferase Reporter Assay kit (Promega, America) under the instruction of the manufacturers 2–3 days after leaf injection.

Electrophoretic Mobility Shift Assay

The coding sequences of *EjRAV1* and *EjRAV2* were cloned into pET28a using primers listed in **Supplementary Table 1**. The 6 × His-EjRAV recombinant proteins and 6 × His control were expressed in *Escherichia coli* BL21 (DE3) cells and induced at OD_{600} = 0.6∼0.8 by 1 mM isopropyl-beta-D-thiogalactoside (IPTG), and oscillated for 4 h at 220 rpm (revolutions per minute) after addition of IPTG. The raw induced proteins were then purified with Amylose Magnetic Beads (NEB, United States). Electrophoretic Mobility Shift Assay (EMSA) assays were performed as previously described (Yu et al., 2020). An EMSA Probe Biotin Labeling Kit (Beyotime, China) was then used to label biotin to the ends of oligonucleotide probes. Excessive unlabeled wild type or mutant probes were used for competition experiments. The probe sequences were listed in **Supplementary Table 1**. The EMSA assay was performed with an EMSA/Gel Shift kit and detected with a chemiluminescent EMSA kit (Beyotime, China). Ultimately, photos of all the gel shift experiments were taken with a ChimeScope series 3300 mini system (Clinxscience, China).

Yeast One-Hybrid Assays

Yeast one-hybrid (Y1H) assays were performed according to the Matchmaker Gold Y1H system user manual (Clontech, America) to detect the binding capacities of *EjRAV1/2* to the *EjFT* promoters. Both *EjFT1* and *EjFT2* promoters contain five putative RAV binding elements (CAACA motif or CACCTG motif, see in **Supplementary Figure 2** and **Supplementary Table 2**). According to **Supplementary Figure 2**, four promoter segments containing all the binding motifs were designed for yeast binding assays, and the segment regions were separately inserted into the pAbAi reporter vector to obtain baits. Meanwhile, the CDS of *EjRAV1/2* was fused to the pGADT7 vector. Then, the recombinant pGADT7-EjRAV1 or pGADT7-EjRAV2 was transformed into the Y1HGold yeast strain with linearized reporter plasmid pAbAi-proFT1 or pAbAi-proFT2 to further determine the protein–DNA interactions. The primers used in these assays were listed in **Supplementary Table 1**.

[4]https://www.dna.affrc.go.jp/PLACE/?action=newplace

RESULTS

Identification of *RAV* Family Genes

A total of six *RAV* family genes were identified in the loquat genome (**Figure 1A**). These genes were clustered into three clades, among which Ej00070074 and Ej00062569 were clustered with four *Arabidopsis RAV*/*TEM* genes (*RAV1*, *RAV1-LIKE*, *RAV2* and *RAV2-LIKE*), Ej00055913, Ej00028520 and Ej00016351 grouped with *FaRAV5*, while Ej00018657 clustered into the *RAV3* clade. These genes showed different expression patterns. Ej00070074 and Ej00062569 (formerly named as *EjRAV1* and *EjRAV2*) were found to be highly expressed in most of the vegetative tissues while the expression levels of the other four genes were scarcely detected (**Figure 1B**). The highly expressed *EjRAV1* and *EjRAV2* were speculated to possibly play a role in repressing flowering or repressing flower bud differentiation in loquat.

EjRAV1/2 Abundance Negatively Associated With the Annual Flowering Pattern

To understand the potential roles of *EjRAV1* and *EjRAV2*, we first analyzed their expression patterns in fruit producing trees of Zaozhong-6 during inflorescence development. We started observations in May because fruits ripened in April and the SAM was still at vegetative growth stage. At this moment, the expression of *EjAP1*, a marker gene for floral initiation, was almost undetectable (**Figures 2A,B**). The SAM of trees stopped growing in July as the vegetative growth ceased, and the floral initiation and inflorescence development started. Ultimately, the SAM was transformed into inflorescence from shoot as the transcription levels of *EjAP1* started to elevate dramatically from 18th July (**Figures 2A,B**).

Abundant *EjRAV1* and *EjRAV2* transcripts were detected at the vegetative growth stage. When the vegetative bud growth ceased, their expression levels were dramatically decreased and maintained at extremely low levels till the floral initiation started (**Figures 2C,D**). In contrast, the expressions levels of *EjFT1* and *EjFT2* increased as *EjRAV1* and *EjRAV2* were down-regulated (**Figure 2E**). Interestingly, the expression levels of both *EjSOC1-1* and *EjSOC1-2* increased after the up-regulation of *EjFT1* and *EjFT2*, and a remarkable increase in the expression levels of *EjAP1* was also detected (**Figure 2F**). The gene expression assays suggested that *EjRAV1* and *EjRAV2* may act as negative regulatory factors in the initiation of flowering in a fruit producing loquat tree.

Abundant *EjRAV1/2* and Rare *EjFT* and *EjSOC1* Transcription Levels in Seedlings at Juvenile Phase

Prolonged juvenile phase is one of the most important ways that guarantee species survival. However, it reduces breeding

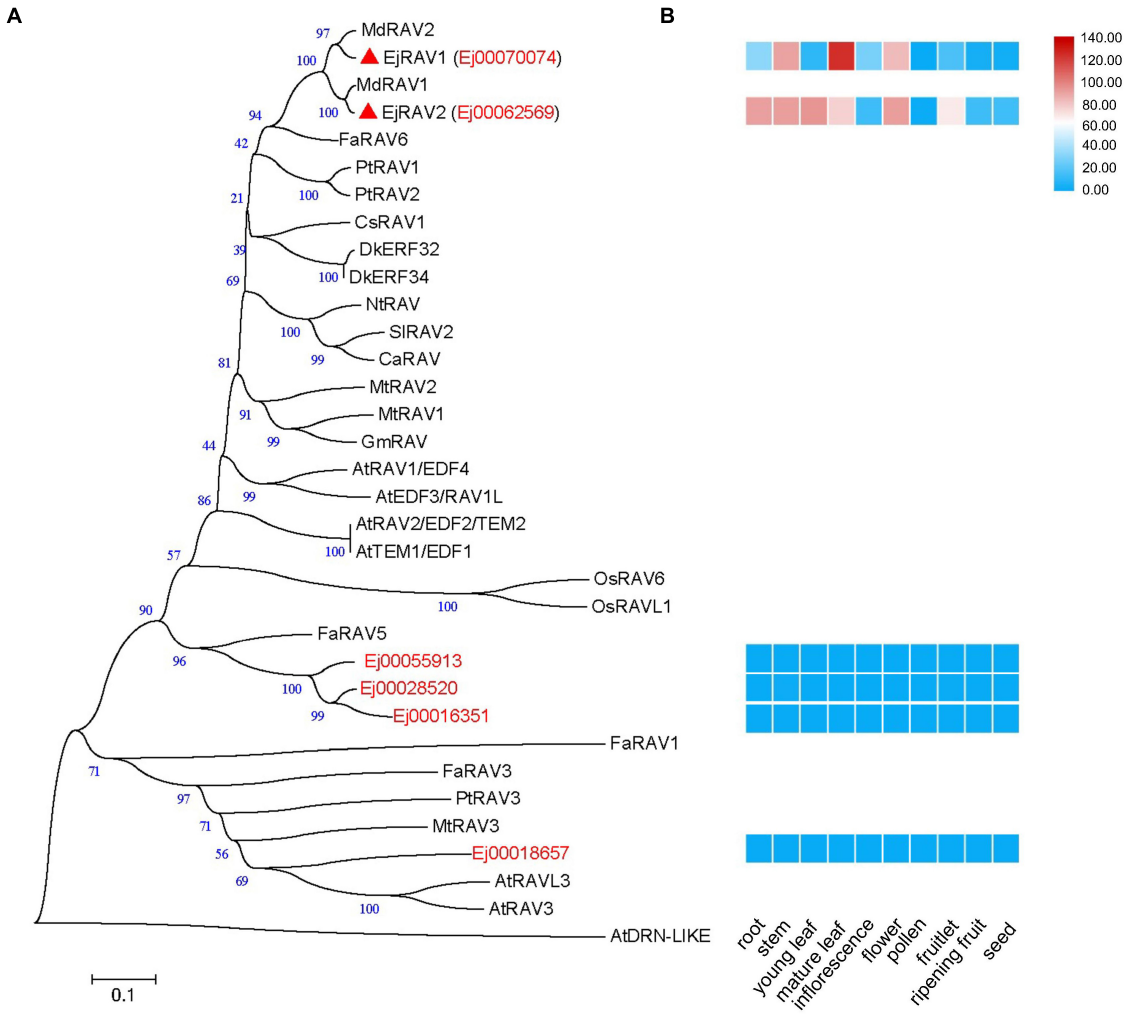


FIGURE 1 | A phylogenetic tree and expression heat map of *RAV* family genes in loquat. **(A)** A phylogenetic tree containing *RAV* homologs identified in loquat genome and other species. AtDRN-LIKE was selected as an outer-group. The red triangles indicate candidate *RAVs* that play role in the transition of reproductive phase in loquat. At, *Arabidopsis thaliana*; Ca, *Capsicum annuum*; Cs, *Castanea sativa*; Dk, *Diospyros kaki*; Ej, *Eriobotrya japonica*; Fa, *Fragaria ananassa*; Gm, *Glycine max*; Md, *Malus domestica*; Mt, *Medicago truncatula*; Nt, *Nicotiana tabacum*; Os, *Oryza sativa*; Pt, *Populus trichocarpa*; Sl, *Solanum lycopersicum*. **(B)** Tissue expression patterns of *EjRAV* family genes based on FPKM values from 10 tissues.

efficiency, especially for perennial crops. To further confirm whether *EjRAV1* and *EjRAV2* play roles in the regulation of juvenile phase in loquat, we analyzed the expression patterns of *EjRAV1/2* and other important flowering regulating genes in 1-year-old seedlings, 2-year-old seedlings, 2-year-old grafted trees, and fruit-bearing trees (**Figure 3A**). The 1-year-old seedlings and 2-year-old seedlings were still at the juvenile phase. The SAMs of both 2-year-old grafted trees (the scions were cut from fruit-bearing trees) and fruit-bearing trees were transformed into inflorescence in the autumn and set fruits later. The abundance of *EjRAV1* and *EjRAV2* transcripts were gradually reduced as the tree achieved the ability to form floral bud step by step (**Figures 3B,C**). By contrast, the expression levels of *EjFT1*, *EjSOC1-1* and *EjSOC1-2* increased continuously over this period (**Figures 3D–F**). Overall, the gene expression analyses suggested that *EjRAV1* and *EjRAV2* may act as important regulators of juvenility phase by repressing the flowering activator gene expressions.

Overexpression of *EjRAV1/2* Prolonged the Juvenile Phase in *Arabidopsis*

To confirm the above-mentioned speculation on the functions of *EjRAV1* and *EjRAV2* in the regulation of flowering time, we cloned each of them into a pGreen-35S plant overexpression vector for overexpression (OE) in *Arabidopsis*. Sequence alignment revealed that the encoded proteins of both genes are characterized by an AP2 domain and a B3 domain, as well as by a nuclear localization signal (NLS) and a B3 repression domain (**Supplementary Figure 3**). OE of both *EjRAV1* and *EjRAV2* in Columbia-0 (Col-0) background significantly delayed flowering (**Figure 4A**). The days to bolting after germination strikingly

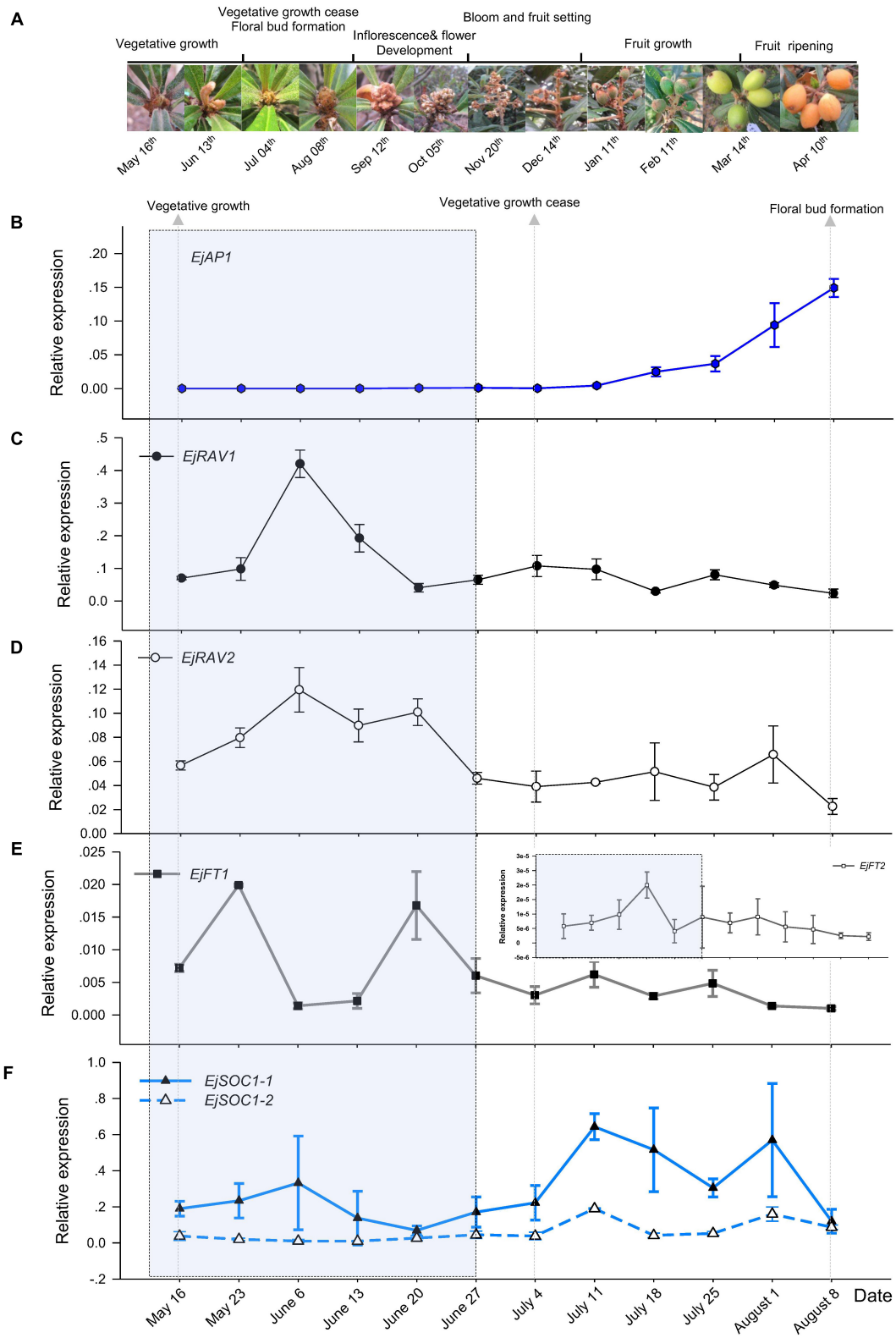


FIGURE 2 | Floral initiation and expressions of flowering regulating gene homologs in "Zaozhong-6" loquat. **(A)** Developmental changes in shoot apical meristem (SAM) of "Zaozhong-6" loquat from May 2018 to April 2019. **(B–F)** Expression patterns of *EjAP1* **(B)**, *EjRAV1* **(C)**, *EjRAV2* **(D)**, *EjFT1/2* **(E)**, and *EjSOC1-1/2* **(F)** during floral initiation and early inflorescence development. *EjAP1* was detected in SAM, while other genes were detected in mature leaves. The error bars represent standard deviation for three biological replicates. The light gray colored window indicates vegetative phase.

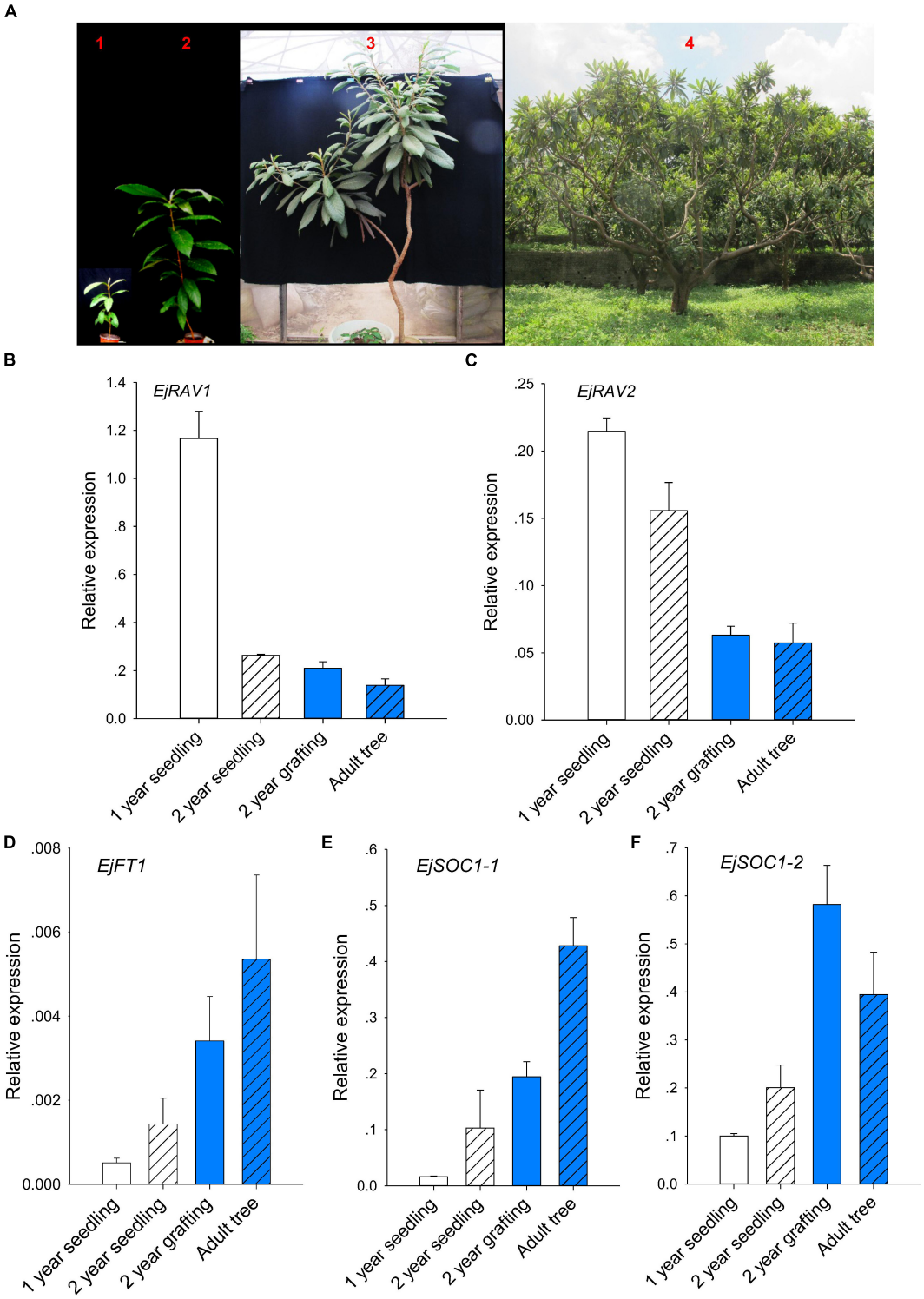


FIGURE 3 | Association of abundant *EjRAVs* expressions and rare flowering activator gene expressions with the juvenile phase. **(A)** "Zaozhong-6" plants at different developmental stages in July, 2018. The red numbers, 1–4, in *A* showed plants of a 1-year-old seedling, a 2-year-old seedling, a 2-year-old grafted tree and an adult fruit-bearing tree used for gene expression assays, respectively. (**B–F**) Expression patterns of *EjRAV1* **(B)**, *EjRAV2* **(C)**, *EjFT1* **(D)**, *EjSOC1-1* **(E)**, and *EjSOC1-2* **(F)** in mature leaves of these trees. The expression level of *EjFT2* was almost undetectable in all the trees and data were not shown. The error bars in **(B–F)** represent standard deviation for three biological replicates.

increased (3–4-fold) for all the OE transgenic plants than the wild type plants (**Figure 4B**), which indicated that *EjRAV1* and *EjRAV2* prolonged the juvenile phase in transgenic plants. At the same time, the number of rosette leaves of each OE transgenic line also greatly increased (**Figure 4C**). Unexpectedly, all the OE transgenic plants had a larger number of branches (about twofold) (**Figures 4D,E**) and stronger root systems than that of Col-0 (**Figure 4E**). It was also interesting that abundant anthocyanins were enriched in the basal parts of petiole and stem for all the OE transgenic plants (**Figure 4F**). Gene expression quantification confirmed that the expression levels of flowering activating genes, including *AtFT*, *AtSOC1*, *AtAP1,* and *AtLFY*, were significantly reduced in EjRAV1-OE and EjRAV2-OE transgenic plants compared to the wild type plants (**Figure 4G**). In addition, *EjRAV1* and *EjRAV2* were overexpressed in the *tem2* mutant. Only stable EjRAV2-OE lines were obtained on the *tem2* background. EjRAV2-OE/*tem2* rescued the phenotype of *tem2* to that of Col-0 wild type plant (**Figure 4H**). The days to bolting and numbers of rosette leaves of EjRAV2-OE/*tem2* transgenic plants were strikingly larger (> 2 fold) than that of the *tem2* mutant (**Figures 4I,J**). In summary, the OE experiments conducted here indicated that *EjRAV1* and *EjRAV2* repress the flowering and activate branching and anthocyanin accumulation in transgenic plants.

EjRAV1/2 Localized in Cell Nuclei and Bind to CAACA Motifs to Repress the Expressions of *EjFT* and *EjSOC1*

In agreement with the conserved NLS) appeared at the 5' terminals of the encoded amino acid sequences (**Supplementary Figure 3**), EjRAV1-GFP and EjRAV2-GFP were localized in the cell nuclei of tobacco leaf cells (**Figure 5A**). To understand whether these two proteins play a role in *EjFT* or *EjSOC1* expression modulation, dual-luciferase assays were performed. A total of 5, 5, 5, and 1 candidate RAV family protein binding sites were identified in the promoter regions of *EjFT1*, *EjFT2*, *EjSOC1-1* and *EjSOC1-2*, respectively (**Supplementary Figure 2** and **Supplementary Table 2**). The luciferase assay data showed that EjRAV1 greatly reduced the expressions of *EjFT1/2* and *EjSOC1-1/2*, while EjRAV2 could significantly inhibit the expressions of *EjFT2* and *EjSOC1-2* (**Figure 5B**). Overall, these results suggested that the *EjRAV1/2* proteins localized in the cell nucleus play regulatory roles in loquat.

To further check the abilities of EjRAV1 and/or EjRAV2 to regulate the expression levels of these genes, EjRAV1 and EjRAV2 proteins were prepared with *Escherichia coli* cells. Specific bands were induced by IPTG with protein size between 40 and 55 KDa and purified with Amylose Magnetic Beads. Protein-DNA binding was then assayed for EjRAV1-*proEjFT1/2* and EjRAV2-*proEjFT1/2* (**Supplementary Figure 4**), EjRAV1-*proEjSOC1-1/2* and EjRAV2-*proEjSOC1-1/2* (**Supplementary Figure 5**) combinations. The EMSA assays indicated that EjRAV1, but not EjRAV2, can bind directly to the CAACA motif at P1 (−1583 to −1579 bp) of *proEjFT1,* but not *proEjFT2, in vitro* (**Figures 5C,D** and **Supplementary Figure 4**). Yeast one-hybrid assay was performed to identify other binding sites on *proEjFT2*, which showed that EjRAV1 and EjRAV2 can directly bind *in vivo* to segment3 (−1118 to −994 bp) that contains two CAACA motifs (**Figure 5E**).

DISCUSSION

Loquat, a Distinct Evergreen Fruit Tree, Blooms in Autumn

The decision to flower in plants during both the early life cycle and annual growth season is a complicated process that ultimately affects reproductive ability, breeding efficiency as well as other research progresses on the species (Somerville and Koornneef, 2002). In addition, flowering time determines the season for fruit, cut-flower and other horticultural products supplies (Wilkie et al., 2008). Loquat is one of the minor fruit trees whose fruits ripen from spring to early summer to meet the market demand of fresh fruits in spring, which benefits the growers greatly (Badenes et al., 2009). Commonly, SAM of perennial plants such as poplar (Tylewicz et al., 2018) and deciduous fruit trees (Kurokura et al., 2013) would cease growth and enter dormancy in fall to enable survival in the cold-drought stressful winter. The time of dormancy release is closely related to the flowering time of deciduous fruit trees including apple, pear, peach, and apricot (Falavigna et al., 2019). Here, we found that loquat initiated floral bud formation (**Figure 2A**) in a similar season compared with that of apple and pear in summer (Wilkie et al., 2008); however, the loquat SAM developed into an indeterminate panicle and bloomed immediately after floral initiation (**Figure 2A**) without the growth cease of reproductive buds or entering dormancy in fall. These results were in accordance with previous observations on the morphological transition of SAM, which also revealed immediate floral initiation in loquat (Liu et al., 2013; Jiang et al., 2019). Collectively, the results from the present and previous studies revealed that the constant development of inflorescence and flower bud without entering dormancy leads to the unique flowering habit (in autumn) of cultivated loquat. Future in-depth studies are required to elucidate how cultivated loquat has evolved this distinct flowering ability.

RAV Homologs Maintain Juvenility and Repress Floral Induction in Loquat

The peculiar flowering habit of loquat ensures the availability of its fresh fruits in spring, a season short of fresh fruits in the market. Lots of endeavors aiming to identify pathways modulating flower induction in loquat have been carried out in the last decade (Esumi et al., 2005; Yang et al., 2007; Fernández et al., 2010; Liu et al., 2013; Zhang et al., 2016; Reig et al., 2017; Jiang et al., 2019, 2020a,b; Chen W. et al., 2020; Xia et al., 2020). *FT* and *SOC1* are floral integrators that link upstream photoperiod and other environmental or internal signals to promote flowering, while the downstream *AP1* controls the identity of floral meristem during flower bud transition (Blümel et al., 2015; Bouché et al., 2016). In loquat, using orthologs of these genes as biomarkers, the developmental

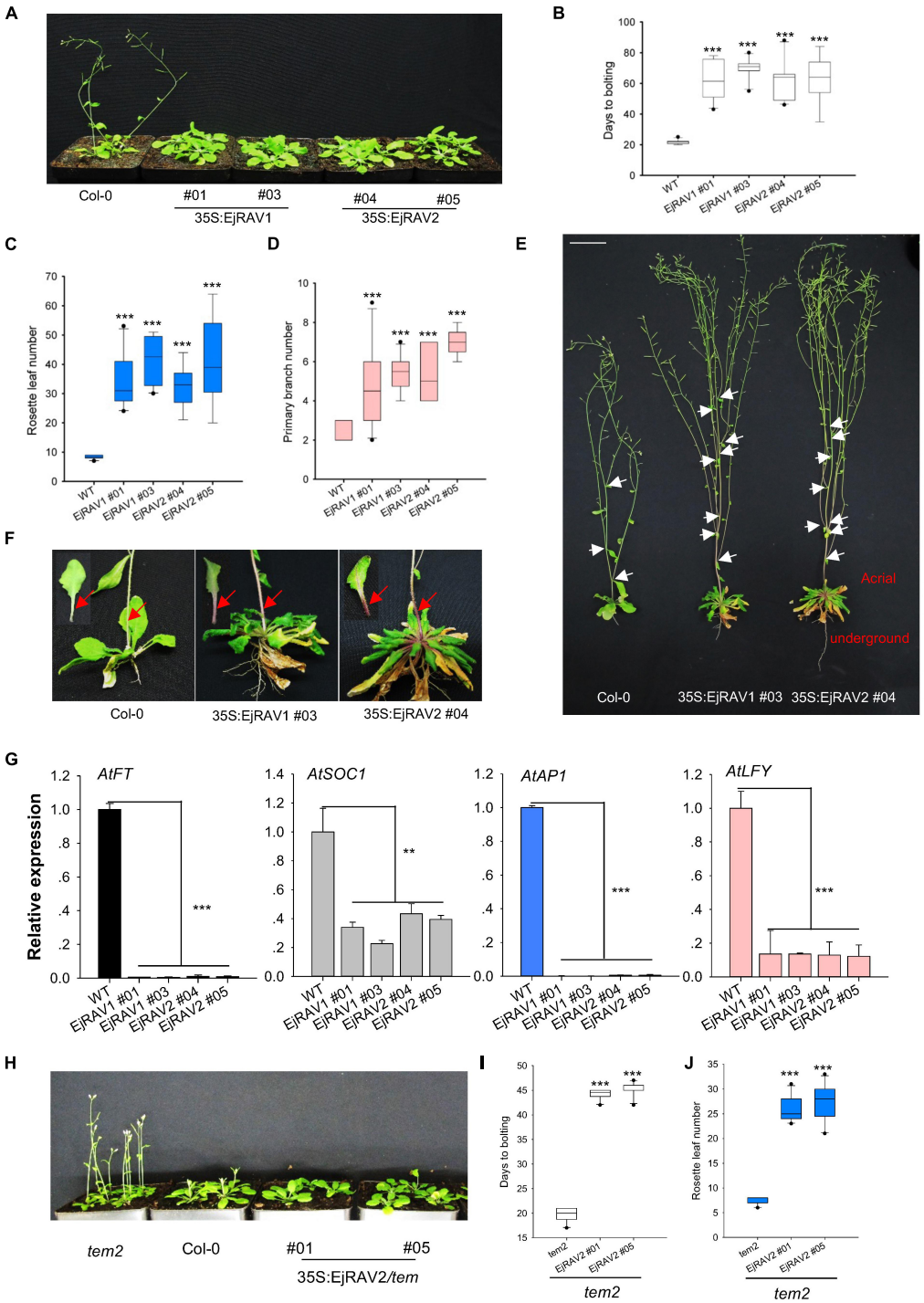


FIGURE 4 | Overexpression of *EjRAV1* and *EjRAV2* delay flowering and promote stem branching in *Arabidopsis thaliana*. **(A)** *EjRAV1*-OE and *EjRAV2*-OE delay flowering in Col-0 ecotype *Arabidopsis*. **(B)** The juvenile length (days to bolting) of all the overexpressed lines were strikingly prolonged. **(C)** The numbers of rosette leaves were greatly increased in the *EjRAV1/2*-OE lines. **(D)** *EjRAV1*-OE and *EjRAV2*-OE promoted lateral branching. **(E)** The phenotypes of Col-0 and *EjRAV1/2*-OE transgenic *Arabidopsis* plants. The ectopic expression of *EjRAV1/2* was not only involved in the regulation of plant architecture (see leaf floral and axis branch numbers) but also associated with strong root system. Arrows indicate the growing point of the primary branch. **(F)** *EjRAV1/2*-OE activated anthocyanin accumulation in petiole and stem. Arrows represent abundant anthocyanin enriched in the basal parts of petiole and stem. **(G)** *EjRAV1*-OE and *EjRAV2*-OE repressed the expressions of flower inducer genes (*FT*, *SOC1*, *AP1*, and *LFY*) in *Arabidopsis*. **(H)** *EjRAV2*-OE rescues the phenotype of *tem2* mutant (Col-0 background) to Col-0 at 4 weeks post germination. The **(I)** juvenile lengths and **(J)** rosette leaf numbers of the *EjRAV2*-OE transgenic plants (on *tem2* background). The ** and *** indicate significant differences at $P < 0.01$ and < 0.001. The error bars in **(B–D,I,J)** represent standard deviation for 15 T_3 plants. The error bars in **(G)** represent standard deviation for three replicates.

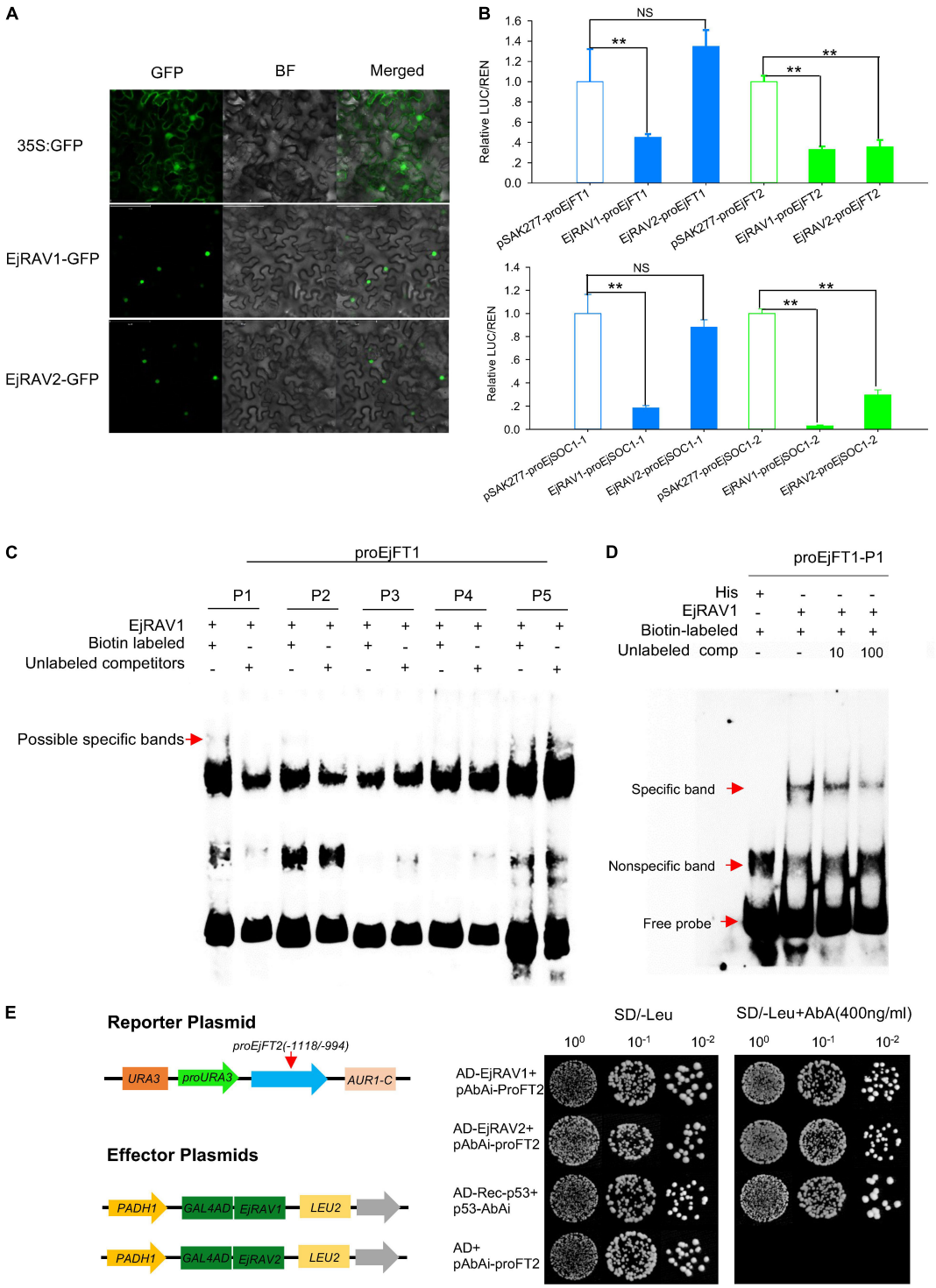


FIGURE 5 | Protein localization, binding and regulation abilities of EjRAVs on FT and SOC1 homologs in loquat. **(A)** Subcellular localization of *EjRAV1/2* proteins expressed in tobacco epidermal cells. **(B)** *EjRAV1/2* regulation on *EjFT*s and *EjSOC1*s expressions in *N. benthamiana* leaves by using the dual-luciferase (LUC) system. Relative renilla (REN) luciferase activity was used as an internal control. Data represent means ± standard error from three replicates. $^{**}P < 0.01$ and NS indicate significance and non-significance, respectively. **(C,D)** Binding ability of EjRAV1 to P1 of *EjFT1* promoter in EMSA assays. P1, P2, P3, P4, and P5 are presumptive RAV protein binding sites on promoter of *EjFT1* at (−1583 to −1579 bp), (−1059 to −1054 bp), (−1033 to −1029 bp), (−782 to −778 bp), and (−73 to −66 bp). Unlabeled probe could reduce binding ability of EjRAV1 to P1 in **(D)**. **(E)** Binding ability of EjRAV1 and EjRAV2 to segment3 (−1118 to −994 bp) of *EjFT2* promoter in Y1H assays.

stages of SAM during the annual growth cycle were classified as vegetative growth, vegetative growth cease and floral bud initiation (**Figure 2**), in accordance with our former histological observations (Jiang et al., 2019). In reference to these stages, negative correlations were detected between the expressions of *EjRAV1/2* and the flowering activators, *FT*, *SOC1,* and *AP1*. Notably, the elevation of expressions of flowering activators occurred after the down-regulation of *EjRAV1/2* (**Figures 2B–F**). These results implied that *EjRAV1/2* are repressors of flowering during the vegetative phase of an annual growth season.

The long juvenile phase would severely reduce breeding efficiency of a perennial crop. Therefore, identifying targets to promote flower initiation at early stage became utmost important (Watson et al., 2018). Although a few flowering regulating genes have been reported to shorten the juvenile period (Hsu et al., 2006; Kotoda et al., 2010; Li et al., 2019), few studies were focused on the identification of candidate genes associated with the promotion or release of juvenility (Missiaggia et al., 2005). In the current study, remarkable negative correlations were discovered between the expressions of *EjRAV1/2* and three flowering activator genes in loquat trees at different growth stages. The abundant accumulation of *EjRAV1/2* was related with the low expressions of *EjFT* and *EjSOC1*s at seedling stage, while high expression levels of *EjFT* and *EjSOC1*s and low levels of *EjRAV1/2* were detected in fruit bearing trees. Similarly in olive and antirrhinum, high *DeTEM* or *AmTEM* gene expressions were detected during the juvenile phase (Sgamma et al., 2015). To further validate the functions of *EjRAV1/2*, we generated overexpressed *EjRAV1* and *EjRAV2* transgenic plants of *Arabidopsis* with two different backgrounds (Col-0 ecotype and *tem2*), and the flowering time of all the transgenic plants were delayed while the juvenile period prolonged strikingly with a sharp decline of expression levels of flowering activator genes (**Figure 4G**). Our results were consistent with previous studies on the overexpression of *AtTEM1*/*AtTEM2* (Castillejo and Pelaz, 2008) and *DeTEM*/*AmTEM* (Sgamma et al., 2015). Here, the gene expression analyses and overexpression experiments revealed that *EjRAV1/2* are vital regulators of prolonged juvenility and repress floral induction in loquat.

EjRAVs Bind to the CAACA Motif of Flowering Signal Integrators to Delay Flower Bud Induction

Binding to the CAACA or CACCTG motif is essential for *RAV* genes to link photoperiod and gibberellin pathways to modulate flowering in *Arabidopsis* (Kagaya et al., 1999; Castillejo and Pelaz, 2008; Osnato et al., 2012). The expression of *FT*, modulated by the balance between CO and TEM, is an integrated signal that triggers flowering (Castillejo and Pelaz, 2008). Here, we identified the binding and transcriptional regulation capacities of *EjRAV1/2* on *FT* and *SOC1* homologs. Our results confirmed that cell nuclei localized *EjRAV1/2* bind to the CAACA motif upstream of *EjFT1* and *EjFT2* promoters to repress their expressions. Although SOC1 is also a flowering pathway integrator and *EjRAV1/2* sharply reduced the expression of *EjSOC1*, no binding site was detected on *EjSOC1*'s promoter region by EMSA (**Supplementary Figure 5**). This might be due to that *EjRAV1/2* influence *EjSOC1*'s expression through another way instead of directly binding to the promoter region of EjSOC1, or due to other reasons, which requires future investigations. While AtRAVs and EjRAVs bind to the CAACA motifs upstream of *FT* promoter regions (Castillejo and Pelaz, 2008), OsRAVL1, a RAV homolog from rice, binds to the E-Box (CACCTG) in the promoters of Brassinosteroid (BR) signaling and the biosynthetic genes (Je et al., 2010).

EjRAVs Promote Branching and Anthocyanin Accumulation

Plant architecture is one of the most important factors modulating grain/fruit production (Wang L. et al., 2020). Loquat is particularly robust in apical dominance and deficient of lateral branches, which remarkably limit the potential of fruit yield improvement (Badenes et al., 2009). Interestingly, the overexpression of *EjRAV1/2* remarkably promoted lateral branches in *Arabidopsis*, similar to the role of their close homologs *CsRAV1* (Moreno-Cortés et al., 2012) and *PtRAV1* (Moreno-Cortés et al., 2017) in the induction of sylleptic branches. Transgenic plants overexpressing *RAV1* lost shoot apical dominance in *Arabidopsis*. In rice, *OsRAVL1* (Je et al., 2010) and *OsRAV6* (Xianwei et al., 2015) maintain BR homeostasis to modulate leaf angle and seed size. Meanwhile, the *TEM* genes in *Arabidopsis* can link photoperiod and gibberellin signals to control flowering (Osnato et al., 2012) and jasmonic acid signaling in salt tolerance (Osnato et al., 2020). These studies suggest that *RAVs* may act to reprogram hormone homeostasis and signal transduction to further promote lateral branches. In *Arabidopsis*, changes on strigolactone signaling affected shoot branching and anthocyanin biosynthesis *via* binding to *BRANCHED1* (*BRC1*) and *PRODUCTION OF ANTHOCYANIN PIGMENT 1* (*PAP1*) promoters (Wang L. et al., 2020). Consistent with this speculation, research in strawberry revealed that *FaRAV1* can positively regulate anthocyanin accumulation in the fruit by activating *FaMYB10*, a homolog of *PAP1* (Zhang et al., 2020). Meanwhile, the vital branching regulators including *BRC1* (Niwa et al., 2013) and it homologs *TEOSINTE BRANCHED 1*, *CYCLOIDEA*, *PCF/TCP* transcription factors (Lucero et al., 2017) were involved in flowering time regulation *via* interacting with *FT* or direct regulation of *SOC1* transcription. Therefore, the *RAV* gene family play vital roles in many developmental processes.

Here, we identified two multifunctional *RAV* homologs in loquat with major roles in maintaining juvenility and delaying flowering in an annual growth cycle. Based on these results, we presented a potential model of pathways of *EjRAV1/2* to modulate floral bud formation during the tree life cycle and annual growth cycle based on molecular assays (**Figure 6**). In this putative model, *EjRAV1/2* bind directly to the CAACA motifs upstream of *EjFT1* and *EjFT2* promoters and repress their expressions to keep loquat trees in the juvenile phase and delay flowering in an annual growth cycle. In addition, *EjRAV1* and *EjRAV2* are important regulators of shoot branching and anthocyanin biosynthesis. The substantial evidence revealed that *EjRAV1/2* play vital roles in the

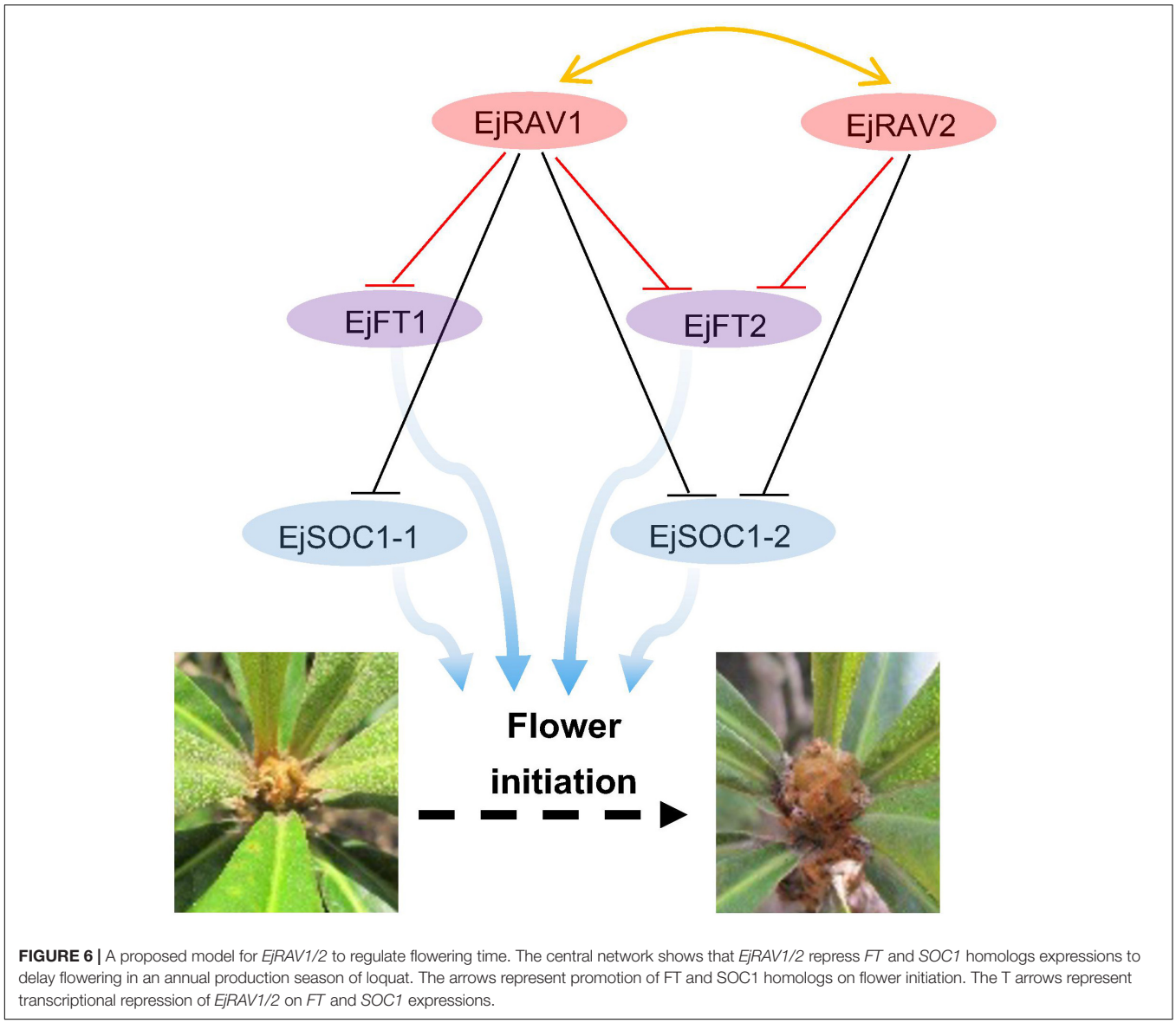


FIGURE 6 | A proposed model for *EjRAV1/2* to regulate flowering time. The central network shows that *EjRAV1/2* repress *FT* and *SOC1* homologs expressions to delay flowering in an annual production season of loquat. The arrows represent promotion of FT and SOC1 homologs on flower initiation. The T arrows represent transcriptional repression of *EjRAV1/2* on *FT* and *SOC1* expressions.

initiation of flowering, and both could be potential targets for breeding programs to accelerate loquat breeding by shortening the juvenile phase and extend the fruit supply season.

CONCLUSION

Two *RELATED TO ABI3 AND VP1 (RAV)* homologs were isolated in loquat. These two genes maintained high expressions during the juvenile or vegetative phase, which was negatively correlated with flowering promoting genes like *EjFT1*/*EjFT2* and *EjSOC1-1*/*EjSOC1-2*. Overexpression of *EjRAV1* and *EjRAV2* delayed flowering and repressed *FT*, *SOC1* and *AP1* expressions in *Arabidopsis*. Both EjRAV1 and EjRAV2 proteins are localized to the cell nuclei and could bind to the CAACA motifs and repress the expressions of *EjFTs* and *EjSOC1s*. These findings contribute to our understanding of the molecular mechanism of loquat phase change and flowering time regulation and provide gene targets to shorten the fruit tree juvenility and accelerate breeding.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

XY, WS, and YG designed the research and obtained the funding. MW and WS performed the transgenic transformation and Y1H assays. MW, WS, and YJ performed the luciferases reporter assays. LZ and WS prepared the proteins and performed the EMSA assays. MW, ZP, WS, CZ, YB, JH, JP, and YG performed

the other experiments. WS, MW, and ZP analyzed the data. ZP, WS, XY, and MS prepared the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Key Research and Development Program (2019YFD1000200), the Key Areas of Science and Technology Planning Project of Guangdong Province (2018B020202011), the National Natural Science Foundation of China (31901973), the Collaborative Innovation Project from the People's Government of Fujian Province and Chinese Academy of Agricultural Sciences (XTCXGC2021006), and Research Start-up Funding (to ZP) from South China Agricultural University.

ACKNOWLEDGMENTS

We thank the experimental instrument service supported by the Key Laboratory of Innovation and Utilization of Horticultural Crop Resources in South China (Ministry of Agriculture). We wish to acknowledge Profs. Wangjin Lu and Jianye Chen for providing pGreen-0800, pGADT7 and pAbAi vectors. We sincerely thank Prof. Shunquan Lin for the guidance and suggestions to this project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.816086/full#supplementary-material

REFERENCES

Andrés, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nat. Rev. Genet.* 13, 627–639. doi: 10.1038/nrg3291

Badenes, M. L., Lin, S., Yang, X., Liu, C., and Huang, X. (2009). "Loquat (*Eriobotrya* Lindl.)," in *Plant Genetics and Genomics: Crops and Models*, eds K. M. Folta and S. E. Gardiner (New York, NY: Springer), 525–538.

Blümel, M., Dally, N., and Jung, C. (2015). Flowering time regulation in crops—what did we learn from *Arabidopsis*? *Curr. Opin. Biotechnol.* 32, 121–129.

Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C. (2016). FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res.* 44, D1167–D1171. doi: 10.1093/nar/gkv1054

Castillejo, C., and Pelaz, S. (2008). The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering. *Curr. Biol.* 18, 1338–1343. doi: 10.1016/j.cub.2008.07.075

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, W., Wang, P., Wang, D., Shi, M., Xia, Y., He, Q., et al. (2020). EjFRI, FRIGIDA (FRI) ortholog from *Eriobotrya japonica*, delays flowering in *Arabidopsis*. *Int. J. Mol. Sci.* 21:1087. doi: 10.3390/ijms21031087

Daniel, P. W. Y. H. (2015). The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: functional evolution and molecular mechanisms. *Mol. Plant* 8, 983–997. doi: 10.1016/j.molp.2015.01.007

Esumi, T., Tao, R., and Yonemori, K. (2005). Isolation of LEAFY and TERMINAL FLOWER 1 homologues from six fruit tree species in the subfamily Maloideae of the Rosaceae. *Sex. Plant Reproduct.* 17, 277–287.

Falavigna, V. D. S., Guitton, B., Costes, E., and Andrés, F. (2019). I want to (bud) break free: the potential role of DAM and SVP-Like genes in regulating dormancy cycle in temperate fruit trees. *Front. Plant Sci.* 9:1990. doi: 10.3389/fpls.2018.01990

Fernández, M. D., Hueso, J. J., and Cuevas, J. (2010). Water stress integral for successful modification of flowering dates in 'Algerie' loquat. *Irrigat. Sci.* 28, 127–134. doi: 10.1007/s00271-009-0165-0

Goeckeritz, C., and Hollender, C. A. (2021). There is more to flowering than those DAM genes: the biology behind bloom in rosaceous fruit trees. *Curr. Opin. Plant Biol.* 59:101995. doi: 10.1016/j.pbi.2020.101995

Gómez-Soto, D., Ramos-Sánchez, J. M., Alique, D., Conde, D., Triozzi, P. M., Perales, M., et al. (2021). Overexpression of a SOC1-related gene promotes bud break in ecodormant poplars. *Front. Plant Sci.* 12:670497. doi: 10.3389/fpls.2021.670497

Halász, J., Hegedűs, A., Karsai, I., Tósaki, Á, and Szalay, L. (2021). Correspondence between SOC1 genotypes and time of endodormancy break in peach (*Prunus persica* L. Batsch) cultivars. *Agronomy* 11:1298. doi: 10.3390/agronomy11071298

Hellens, R. P., Allan, A. C., Friel, E. N., Bolitho, K., Grafton, K., Templeton, M. D., et al. (2005). Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant Methods* 1:13. doi: 10.1186/1746-4811-1-13

Henriques, R., Lin, S., Chua, N., Zhang, X., and Niu, Q. (2006). *Agrobacterium* - mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat. Protoc.* 1, 641–646. doi: 10.1038/nprot.2006.97

Hsu, C. Y., Liu, Y., Luthe, D. S., and Yuceer, C. (2006). Poplar FT2 shortens the juvenile phase and promotes seasonal flowering. *Plant Cell* 18, 1846–1861. doi: 10.1105/tpc.106.041038

Je, B. I., Piao, H. L., Park, S. J., Park, S. H., Kim, C. M., Xuan, Y. H., et al. (2010). RAV-Like1 maintains brassinosteroid homeostasis *via* the coordinated activation of BRI1 and biosynthetic genes in rice. *Plant Cell* 22, 1777–1791.

Jia, X. L., Chen, Y. K., Xu, X. Z., Shen, F., Zheng, Q. B., Du, Z., et al. (2017). miR156 switches on vegetative phase change under the regulation of redox signals in apple seedlings. *Sci. Rep.* 7, 14213–14223. doi: 10.1038/s41598-017-14671-8

Jiang, Y., Peng, J., Wang, M., Su, W., Gan, X., Jing, Y., et al. (2020a). The role of EjSPL3, EjSPL4, EjSPL5, and EjSPL9 in regulating flowering in loquat (*Eriobotrya japonica* Lindl.). *Int. J. Mol. Sci.* 21:248. doi: 10.3390/ijms21010248

Jiang, Y., Peng, J., Zhu, Y., Su, W., Zhang, L., Jing, Y., et al. (2019). The role of EjSOC1s in flower initiation in *Eriobotrya japonica*. *Front. Plant Sci.* 10:253. doi: 10.3389/fpls.2019.00253

Jiang, Y., Zhu, Y., Zhang, L., Su, W., Peng, J., Yang, X., et al. (2020b). EjTFL1 genes promote growth but inhibit flower bud differentiation in loquat. *Front. Plant Sci.* 11:576. doi: 10.3389/fpls.2020.00576

Kagaya, Y., Ohmiya, K., and Hattori, T. (1999). RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Res.* 27, 470–478. doi: 10.1093/nar/27.2.470

Kotoda, N., Hayashi, H., Suzuki, M., Igarashi, M., Hatsuyama, Y., Kidou, S., et al. (2010). Molecular characterization of FLOWERING LOCUS T-Like genes of apple (*Malus* × *domestica* Borkh.). *Plant Cell Physiol.* 51, 561–575. doi: 10.1093/pcp/pcq021

Kurokura, T., Mimida, N., Battey, N. H., and Hytonen, T. (2013). The regulation of seasonal flowering in the *Rosaceae*. *J. Exp. Bot.* 64, 4131–4141. doi: 10.1093/jxb/ert233

Li, C., Yamagishi, N., Kasajima, I., and Yoshikawa, N. (2019). Virus-induced gene silencing and virus-induced flowering in strawberry (*Fragaria* × *ananassa*) using apple latent spherical virus vectors. *Hortic. Res.* 6:18. doi: 10.1038/s41438-018-0106-2

Liu, Y., Song, H., Liu, Z., Hu, G., and Lin, S. (2013). Molecular characterization of loquat EjAP1 gene in relation to flowering. *Plant Growth Regul.* 70, 287–296. doi: 10.1007/s10725-013-9800-0

Liu, Y., Zhao, Q., Meng, N., Song, H., Li, C., Hu, G., et al. (2017). Over-expression of EjLFY-1 leads to an early flowering habit in strawberry (*Fragaria* × *ananassa*) and its asexual progeny. *Front. Plant Sci.* 8:496. doi: 10.3389/fpls.2017.00496

Lucero, L. E., Manavella, P. A., Gras, D. E., Ariel, F. D., and Gonzalez, D. H. (2017). Class I and class II TCP transcription factors modulate SOC1-dependent flowering at multiple levels. *Mol. Plant* 10, 1571–1574. doi: 10.1016/j.molp.2017.09.001

Matias-Hernandez, L., Aguilar-Jaramillo, A. E., Marin-Gonzalez, E., Suarez-Lopez, P., and Pelaz, S. (2014). RAV genes: regulation of floral induction and beyond. *Ann. Bot.* 114, 1459–1470. doi: 10.1093/aob/mcu069

McGarry, R. C., Klocko, A. L., Pang, M., Strauss, S. H., and Ayre, B. G. (2017). Virus-Induced flowering: an application of reproductive biology to benefit plant research and breeding. *Plant Physiol.* 173, 47–55. doi: 10.1104/pp.16.01336

Missiaggia, A. A., Piacezzi, A. L., and Grattapaglia, D. (2005). Genetic mapping of Eef1, a major effect QTL for early flowering in *Eucalyptus grandis*. *Tree Genet. Genomes* 1, 79–84. doi: 10.1007/s11295-005-0011-3

Moreno-Cortés, A., Hernández-Verdeja, T., Sánchez-Jiménez, P., González-Melendi, P., Aragoncillo, C., and Allona, I. (2012). CsRAV1 induces sylleptic branching in hybrid poplar. *New Phytol.* 194, 83–90. doi: 10.1111/j.1469-8137.2011.04023.x

Moreno-Cortés, A., Ramos-Sánchez, J. M., Hernández-Verdeja, T., González-Melendi, P., Alves, A., Simoes, R., et al. (2017). Impact of RAV1-engineering on poplar biomass production: a short-rotation coppice field trial. *Biotechnol. Biofuels* 10:110. doi: 10.1186/s13068-017-0795-z

Niwa, M., Daimon, Y., Kurotani, K., Higo, A., Pruneda-Paz, J. L., Breton, G., et al. (2013). BRANCHED1 interacts with FLOWERING LOCUS T to repress the floral transition of the axillary meristems in *Arabidopsis*. *Plant Cell* 25, 1228–1242. doi: 10.1105/tpc.112.109090

Osnato, M., Castillejo, C., Matías-Hernández, L., and Pelaz, S. (2012). TEMPRANILLO genes link photoperiod and gibberellin pathways to control flowering in *Arabidopsis*. *Nat. Commun.* 3:808. doi: 10.1038/ncomms1810

Osnato, M., Cereijo, U., Sala, J., Matias-Hernandez, L., Aguilar-Jaramillo, A. E., Rodríguez-Goberna, M. R., et al. (2020). The floral repressors TEMPRANILLO1 and 2 modulate salt tolerance by regulating hormonal components and photoprotection in *Arabidopsis*. *Plant J.* 105, 7–12. doi: 10.1111/tpj.15048

Potter, D., Eriksson, T., Evans, R. C., Oh, S., Smedmark, J. E. E., Morgan, D. R., et al. (2007). Phylogeny and classification of Rosaceae. *Plant Syst. Evol.* 266, 5–43. doi: 10.1007/s00606-007-0539-9

Reig, C., Gil-Muñoz, F., Vera-Sirera, F., García-Lorca, A., Martínez-Fuentes, A., Mesejo, C., et al. (2017). Bud sprouting and floral induction and expression of FT in loquat [*Eriobotrya japonica* (Thunb.) Lindl.]. *Planta* 246, 915–925. doi: 10.1007/s00425-017-2740-6

Sgamma, T., Jackson, A., Muleo, R., Thomas, B., and Massiah, A. (2015). TEMPRANILLO is a regulator of juvenility in plants. *Sci. Rep.* 4:3704.

Somerville, C., and Koornneef, M. (2002). A fortunate choice: the history of *Arabidopsis* as a model plant. *Nat. Rev. Genet.* 3, 883–889.

Song, Y. H., Ito, S., and Imaizumi, T. (2013). Flowering time regulation: photoperiod- and temperature-sensing in leaves. *Trends Plant Sci.* 18, 575–583. doi: 10.1016/j.tplants.2013.05.003

Su, W., Jing, Y., Lin, S., Yue, Z., Yang, X., Xu, J., et al. (2021). Polyploidy underlies co-option and diversification of biosynthetic triterpene pathways in the apple tribe. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2101767118. doi: 10.1073/pnas.2101767118

Su, W., Yuan, Y., Zhang, L., Jiang, Y., Gan, X., Bai, Y., et al. (2019). Selection of the optimal reference genes for expression analyses in different materials of *Eriobotrya japonica*. *Plant Methods* 15:7. doi: 10.1186/s13007-019-0391-2

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., et al. (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

Tränkner, C., Lehmann, S., Hoenicka, H., Hanke, M. V., Fladung, M., Lenhardt, D., et al. (2010). Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. *Planta* 232, 1309–1324. doi: 10.1007/s00425-010-1254-2

Tylewicz, S., Petterle, A., Marttila, S., Miskolczi, P., Azeez, A., Singh, R. K., et al. (2018). Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* 360, 212–215. doi: 10.1126/science.aan8576

Voogd, C., Brian, L. A., Wang, T., Allan, A. C., and Varkonyi-Gasic, E. (2017). Three FT and multiple CEN and BFT genes regulate maturity, flowering, and vegetative phenology in kiwifruit. *J. Exp. Bot.* 68, 1539–1553.

Wang, J. (2014). Regulation of flowering time by the miR156-mediated age pathway. *J. Exp. Bot.* 65, 4723–4730. doi: 10.1093/jxb/eru246

Wang, J., Gao, Z., Li, H., Jiu, S., Qu, Y., Wang, L., et al. (2020). Dormancy-associated MADS-Box (DAM) genes influence chilling requirement of sweet cherries and co-regulate flower development with SOC1 gene. *Int. J. Mol. Sci.* 21:921. doi: 10.3390/ijms21030921

Wang, L., Wang, B., Yu, H., Guo, H., Lin, T., Kou, L., et al. (2020). Transcriptional regulation of strigolactone signalling in *Arabidopsis*. *Nature* 583, 277–281. doi: 10.1038/s41586-020-2382-x

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M. D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8

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

Xia, Y., Xue, B., Shi, M., Zhan, F., Wu, D., Jing, D., et al. (2020). Comparative transcriptome analysis of flower bud transition and functional characterization of EjAGL17 involved in regulating floral initiation in loquat. *PLoS One* 15:e239382. doi: 10.1371/journal.pone.0239382

Xiang, Y., Huang, C., Hu, Y., Wen, J., Li, S., Yi, T., et al. (2016). Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. *Mol. Biol. Evol.* 34, 262–281. doi: 10.1093/molbev/msw242

Xianwei, S., Zhang, X., Sun, J., and Cao, X. (2015). Epigenetic mutation of RAV6 affects leaf angle and seed size in rice. *Plant Physiol.* 169, 2118–2128. doi: 10.1104/pp.15.00836

Xu, M., Hu, T., Zhao, J., Park, M., Earley, K. W., Wu, G., et al. (2016). Developmental functions of miR156-regulated SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes in *Arabidopsis thaliana*. *PLoS Genet.* 12:e1006263. doi: 10.1371/journal.pgen.1006263

Xu, Y., Zhang, L., and Wu, G. (2018). Epigenetic regulation of juvenile-to-adult transition in plants. *Front. Plant Sci.* 9:1048. doi: 10.3389/fpls.2018.01048

Yamagishi, N., Li, C., and Yoshikawa, N. (2016). Promotion of flowering by apple latent spherical virus vector and virus elimination at high temperature allow accelerated breeding of apple and pear. *Front. Plant Sci.* 7:171.

Yang, S., Hu, Y., and Zhou, B. (2007). Studies on floral induction phase and carbohydrate contents in apical bud during floral differentiation in loquat. *Acta Hortic.* 750, 395–400.

Yarur, A., Soto, E., León, G., and Almeida, A. M. (2016). The sweet cherry (*Prunus avium*) FLOWERING LOCUS T gene is expressed during floral bud determination and can promote flowering in a winter-annual *Arabidopsis* accession. *Plant Reproduct.* 29, 311–322. doi: 10.1007/s00497-016-0296-4

Yu, H., Zhang, L., Wang, W., Tian, P., Wang, W., Wang, K., et al. (2020). TCP5 controls leaf margin development by regulating the KNOX and BEL-like transcription factors in *Arabidopsis*. *J. Exp. Bot.* 72, 1809–1821. doi: 10.1093/jxb/eraa569

Zhang, L., Jiang, Y., Zhu, Y., Su, W., Long, T., Huang, T., et al. (2019). Functional characterization of GI and CO homologs from *Eriobotrya deflexa* Nakai forma koshunensis. *Plant Cell Rep.* 38, 533–543. doi: 10.1007/s00299-019-02384-3

Zhang, L., Yu, H., Lin, S., and Gao, Y. (2016). Molecular characterization of FT and FD homologs from *Eriobotrya deflexa* Nakai forma koshunensis. *Front. Plant Sci.* 7:8. doi: 10.3389/fpls.2016.00008

Zhang, Z., Shi, Y., Ma, Y., Yang, X., Yin, X., Zhang, Y., et al. (2020). The strawberry transcription factor FaRAV1 positively regulates anthocyanin accumulation by activation of FaMYB10 and anthocyanin pathway genes. *Plant Biotechnol. J.* 18, 2267–2279. doi: 10.1111/pbi.13382

Zheng, S., Huang, J., and Xu, X. (1991). Research on early maturing and related properties in loquat. *Southeast Hortic.* 5, 1–5.

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 January 2022
doi: 10.3389/fpls.2021.802016

The LcKNAT1-LcEIL2/3 Regulatory Module Is Involved in Fruitlet Abscission in Litchi

Xingshuai Ma[1,2], Peiyuan Ying[2], Zidi He[2], Hong Wu[2], Jianguo Li[1,2] and Minglei Zhao[1,2,3]**

[1] Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China, [2] State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in South China at Ministry of Agriculture and Rural Affair, Guangdong Litchi Engineering Research Center, College of Horticulture, South China Agricultural University, Guangzhou, China, [3] Guangdong Provincial Key Laboratory of Postharvest Science of Fruits and Vegetables, College of Horticulture, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Yuexue Liu, Shenyang Agricultural University, China

Reviewed by:
Tao Xu, Shenyang Agricultural University, China
Agata Kućko, Warsaw University of Life Sciences, Poland
Francisco Ramón Tadeo, Instituto Valenciano de Investigaciones Agrarias, Spain

***Correspondence:**
Jianguo Li
jianli@scau.edu.cn
Minglei Zhao
zml503@scau.edu.cn

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

Received: *26 October 2021*
Accepted: *15 December 2021*
Published: *21 January 2022*

Citation:
Ma X, Ying P, He Z, Wu H, Li J and Zhao M (2022) The LcKNAT1-LcEIL2/3 Regulatory Module Is Involved in Fruitlet Abscission in Litchi. Front. Plant Sci. 12:802016. doi: 10.3389/fpls.2021.802016

Large and premature organ abscission may limit the industrial development of fruit crops by causing serious economic losses. It is well accepted that ethylene (ET) is a strong inducer of organ abscission in plants. However, the mechanisms underlying the control of organ abscission by ET are largely unknown. We previously revealed that LcKNAT1, a KNOTTED-LIKE FROM *ARABIDOPSIS* THALIANA1 (KNAT1)-like protein, acted as a negative regulator in control of fruitlet abscission through suppressing the expression of ET biosynthetic genes in litchi. In this study, we further reported that LcKNAT1 could also directly repress the transcription of LcEIL2 and LcEIL3, two ETHYLENE INSENSITIVE 3-like (EIL) homologs in litchi, which functioned as positive regulators in ET-activated fruitlet abscission by directly promoting the expression of genes responsible for ET biosynthesis and cell wall degradation. The expression level of *LcKNAT1* was downregulated, while *LcEIL2/3* was upregulated at the abscission zone (AZ) accompanying the fruitlet abscission in litchi. The results of electrophoretic mobility shift assays (EMSAs) and transient expression showed that LcKNAT1 could directly bind to the promoters of *LcEIL2* and *LcEIL3* and repress their expression. Furthermore, the genetic cross demonstrated that the β-glucuronidase (GUS) expression driven by the promoters of *LcEIL2* or *LcEIL3* at the floral AZ was obviously suppressed by LcKNAT1 under stable transformation in *Arabidopsis*. Taken together, our findings suggest that the LcKNAT1-LcEIL2/3 regulatory module is likely involved in the fruitlet abscission in litchi, and we propose that LcKNAT1 could suppress both ET biosynthesis and signaling to regulate litchi fruit abscission.

Keywords: litchi (*Litchi chinensis* Sonn.), fruitlet abscission, LcKNAT1, LcEIL2/3, ethylene

INTRODUCTION

Organ abscission is a widespread phenomenon throughout the life cycle of plants that occurs not only under stressful environmental conditions (biotic and abiotic) but also as a result of its own developmental process (Sexton and Roberts, 1982; Patterson, 2001; Taylor and Whitelaw, 2001; Roberts et al., 2002; Estornell et al., 2013). A high incidence of organ abscission (leaves, flowers, or fruits) may limit yield in crop species and, therefore, make their cultivation economically not

viable. Thus, understanding the molecular mechanism of abscission is of great significance for the development of the crop industry.

Many factors coordinate the abscission process, of which plant hormones play a critical role. It is now well accepted that ethylene (ET) acts as a strong inducer in organ abscission. ET is formed from methionine *via* S-adenosyl-L-methionine (AdoMet) and 1-aminocyclopropane-1-carboxylic acid (ACC). ACC synthase (ACS) and ACC oxidase (ACO) catalyze the conversion of AdoMet to ACC and of ACC to ET, respectively (Yang and Hoffman, 1984). The effect of ET on organ abscission is closely related to its biosynthesis and signal transduction pathway. On the one hand, the increase of ET production was found in organs prior to abscission (Meir et al., 2019). On the other hand, organ abscission was induced or obviously accelerated after the application of exogenous ET, such as fruit abscission in mango (Nunez-Elisea and Davenport, 1986), olive (Kitsaki et al., 1999), apple (Dal et al., 2005), grape berry (Uzquiza et al., 2014), and litchi (Li et al., 2015; Ma et al., 2021), leaf abscission in citrus (Gomez-Cadenas et al., 1996), and flower abscission in *Lupinus luteus* (Wilmowicz et al., 2016). As the core downstream transcription factor (TF) of ET signaling, ETHYLENE INSENSITIVE 3 (EIN3) and EIN3-LIKE (EIL) proteins are considered to play a crucial role in ET-induced organ abscission. For example, *OeEIL2* was upregulated under ET treatment and was identified to be associated with mature fruit abscission in olive (Parra-Lobato and Gomez-Jimenez, 2011). ET treatment accelerated the leaf abscission of Chinese cabbage, in which the expression of *BcEIL1/2/3* was significantly induced (Meng et al., 2019). After knocking down the expression of *LeEILs*, the flower abscission was obviously reduced in tomatoes (Tieman et al., 2001). Recently, we also found that LcEIL2/3 functioned as positive regulators of fruitlet abscission in litchi, probably by activating the expression of genes responsible for ET biosynthesis and cell wall degradation (Ma et al., 2020). Taken together, previous studies strongly demonstrate that organ abscission is closely associated with ET biosynthesis and signaling. However, how the *ACS/ACO* genes are regulated and the mechanism underlying EIN3/EILs in the regulation of abscission remains elusive.

Litchi (*Litchi chinensis* Sonn.) is an important tropical and subtropical economic fruit crop. It is widely cultivated in over 20 countries. However, there are three or four waves of physiological fruit abscission dependent on cultivars throughout fruit development, leading to heavy economic loss and extremely limiting the development of the litchi industry (Yuan and Huang, 1988; Mitra et al., 2003). In *Arabidopsis*, KNOTTED-LIKE FROM *ARABIDOPSIS* THALIANA1/BREVIPEDICELLUS (KNAT1/BP) has been proven to be a key downstream factor of the INFLORESCENCE DEFICIENT IN ABSCISSION-HAESA/HAESA-LIKE2 (IDA-HAE/HSL2) pathway that controls the floral organ abscission in an ET-independent mode (Butenko et al., 2003; Cho et al., 2008; Stenvik et al., 2008; Shi et al., 2011; Santiago et al., 2016). In a previous study, we identified a KNAT1 homolog in litchi, i.e., LcKNAT1; interestingly, LcKNAT1 could suppress the ET production to control the fruitlet abscission by directly binding to the promoter of ET biosynthesis genes (Zhao et al., 2020). In this study, we further reported that LcKNAT1 could also directly bind to the promoters of *LcEIL2/3* and suppress their expression. These results suggest that LcKNAT1 could regulate the fruitlet abscission by suppressing both ET biosynthesis and signaling in litchi.

MATERIALS AND METHODS

Plant Materials and Treatments

Three 20-year-old "Feizixiao" litchi trees (*L. chinensis* Sonn.) were randomly selected in the orchard of the South China Agricultural University. Twenty fruit-bearing shoots about 5–8 mm in different directions were chosen. Treatment of fruitlet removal at 25 days after anthesis was applied in ten shoots, and the other ten were used as control. Among each treatment, three shoots were used to monitor the dynamics of peduncle abscission and the others were used for sampling. Fruitlet abscission zone (FAZ) samples were collected and stored at -80°C for future analyses. Each tree was treated as a biological replicate (Li et al., 2015).

Generation of Transgenic Plants

The coding sequence of *LcKNAT1* was cloned and infused into vector pCAMBIA1302 driven by the CaMV 35S promoter. Then, the recombinant construct was transformed into *Arabidopsis* Col plants according to the floral dip method (Clough and Bent, 1998). Transgenic lines were selected by hygromycin resistance and analyzed by quantitative reverse transcription PCR (qRT-PCR). T3 homozygous transgenic *Arabidopsis* plants were used for further analysis.

2′,7′-Bis-(2-Carboxyethyl)-5-(and-6)-Carboxyfluorescein Fluorescence Analyses

The 2′,7′-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF) fluorescence analysis was carried out to measure cytosolic pH. According to the previously described studies (Sundaresan et al., 2014; Ying et al., 2016), *Arabidopsis* flowers were removed from the plant body and soaked into 10 μM BCECF-AM (B1150, Thermo Scientific) solution for 20 min under darkness. Then, phosphate-buffered saline (PBS, pH 7.4) was used to remove the excess BCECF-AM. Images were taken using the confocal laser scanning microscope (LSM 7 DUO, ZEISS, Germany). BCECF fluorescence and chlorophyll autofluorescence were detected under 488 nm and 633 nm light, respectively.

Histochemical β-Glucuronidase Assays and Genetic Cross Assays

Notably, 1,500 bp promoter sequence of *LcEIL2* or *LcEIL3* was cloned and infused into vector pCAMBIA1391 which contains the β-glucuronidase (GUS) reporter, respectively. The recombinant constructs were transformed into *Arabidopsis* Col plants to obtain T3 homozygous Pro_{LcEIL2}*:GUS* or Pro_{LcEIL3}*:GUS* plants. Then, T3 homozygous Pro_{LcEIL2}*:GUS*

plants or Pro_{LcEIL3}:*GUS* plants were crossed with T3 homozygous *35S:LcKNAT1-2* plants, and F1 *35S:LcKNAT1-2* × Pro_{LcEIL2}:*GUS* plants or *35S:LcKNAT1-2* × Pro_{LcEIL3}:*GUS* plants were used for GUS staining. In brief, flowers/siliques of transgenic plants were incubated in GUS staining buffer for 6 h at 37°C and then decolorized using 70% ethanol. GUS signals were detected by a stereoscope (ZEISS, SV11).

Real-Time Polymerase Chain Reaction Analysis

Total RNA was extracted using the Column Plant RNAout Kit (TIANDZ, Beijing). The TransScript One-Step gDNA Removal Kit and the cDNA Synthesis SuperMix Kit (TransGen, Beijing) were used to generate the first-strand cDNA. qRT-PCR was performed using SYBR Green Polymerase Chain Reaction (PCR) SuperMix (Bio-Rad) on an ABI7500 Real-Time PCR System (Applied Biosystems). The expression level was analyzed using elongation factor 1-alpha (*EF-1a*) and *Ubiquitin 10* (*UBQ*) as an internal reference for litchi and *Arabidopsis*, respectively (Zhong et al., 2011; Ying et al., 2016). Three biological replicates were carried out.

Yeast One-Hybrid Assay

Yeast one-hybrid (Y1H) assay was performed by using the Matchmaker Gold Yeast One-Hybrid System (Clontech). The coding sequence of LcKNAT1 was inserted into the pGADT7 vector. *LcEIL2* or *LcEIL3* promoter was cloned into the pAbAi vector. The recombinant pAbAi constructs were linearized and then transformed into the Y1H Gold strain, respectively. Positive yeast cells were transformed with pGADT7-LcKNAT1. The DNA–protein interaction was assessed based on the growth ability of transformed yeast on SD/-Leu medium with aureobasidin A (AbA).

Electrophoretic Mobility Shift Assay

The sequence containing the DNA-binding domain of LcKNAT1 was cloned into vector PGEX-4T-1, and then, the recombinant construct was transformed into *Escherichia coli* strain BM Rosetta (DE3). The GST-LcKNAT1 fusion protein was purified using GSTSep Glutathione Agarose Resin (YEASEN, Shanghai). The probes including KNOX binding sites were biotin-labeled at the 3' end, deriving from *LcEIL2* and *LcEIL3* promoters. The cold probes, containing the same sequences without biotin-labeled, and the mutant probes were used as competitors. These probes were incubated with GST-LcKNAT1 fusion protein. The DNA-binding assays were detected using the LightShift Chemiluminescent EMSA Kit (Thermo Scientific, Illinois, United States). ChemiDoc MP Imaging System (Bio-Rad, Hercules, CA, United States) was used to take the binding photos.

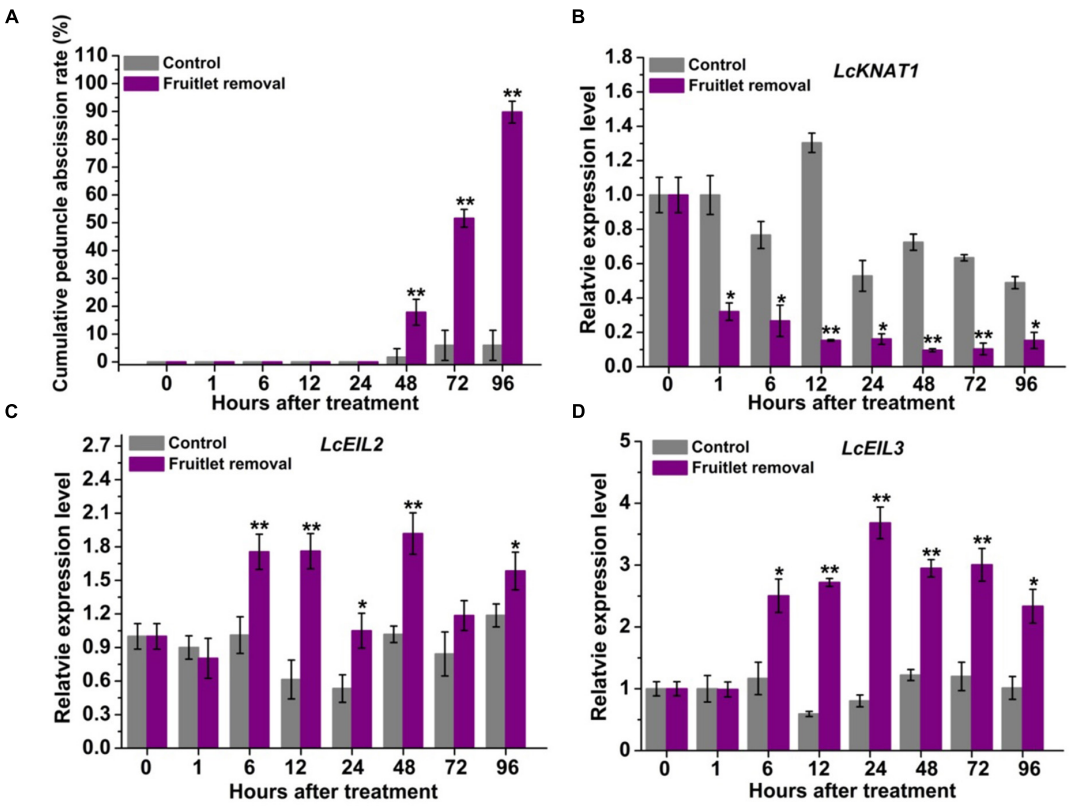


FIGURE 1 | *LcKNAT1* is decreased, but *LcEIL2/3* expression is increased in the abscission zone (AZ) during the fruitlet removal-induced abscission process in litchi. **(A)** Fruitlet removal induced the peduncle abscission. The expression patterns of *LcKNAT1* **(B)**, *LcEIL2* **(C)**, and *LcEIL3* **(D)** after fruitlet removal. SE comes from three replicates. Asterisks indicate a significant difference (Student's *t*-test: $^{*}P < 0.05$, $^{**}P < 0.01$).

Dual-Luciferase Reporter Assay

To detect the effect of LcKNAT1 on the transcriptional ability of *LcEIL2* and *LcEIL3*, a dual-luciferase (LUC) transient expression system was performed. The coding sequence of LcKNAT1 infused into pGreenII 62-SK vector was used as the effector. The promoter sequence of *LcEIL2* or *LcEIL3* inserted into pGreenII 0800-LUC vector was used as the reporter. Each pair of effector and reporter was incubated and co-transformed into tobacco leaves as reported previously (Ma et al., 2019). Notably, 2 days later, the Dual-Luciferase Assay reagents (Promega) were used to detect the LUC and renilla (REN) LUC activity, and then, the ratio of LUC/REN was calculated. For each pair, at least six independent replicates were performed.

Statistical Analysis

Data were reported as the means ± SE. Statistical significance was detected according to the one-way ANOVA using SPSS (version 21.0, Chicago, IL, United States).

RESULTS

LcKNAT1 Is Decreased but *LcEIL2/3* Are Increased at Abscission Zone Accompanying the Fruitlet Abscission Process in Litchi

Previously, *LcKNAT1* was found to be significantly downregulated, while *LcEIL2/3* was obviously upregulated

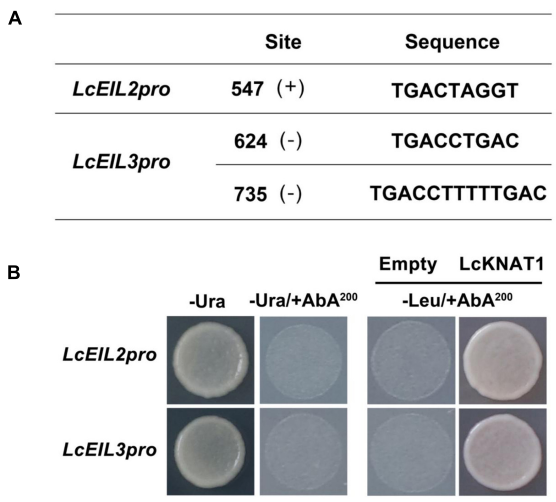


FIGURE 2 | Binding of LcKNAT1 to the promoters of *LcEIL2* and *LcEIL3* by Yeast one-hybrid (Y1H) assay. **(A)** The core TGAC *cis*-elements were responsible for KNOX protein binding in the promoters of *LcEIL2* and *LcEIL3*. **(B)** The Y1H analysis of LcKNAT1 binding to *LcEIL2/3* promoters. Left: No basal activities of LcEIL2 or LcEIL3 promoter were detected in yeast grown on SD medium lacking Ura in the presence of aureobasidin A (AbA). Right: Yeast growth assay after the Y1H reporter strains were co-transformed with LcKNAT1 effector. The interaction was evaluated based on the growth conditions of transformed yeast on SD medium lacking Leu in the presence of AbA.

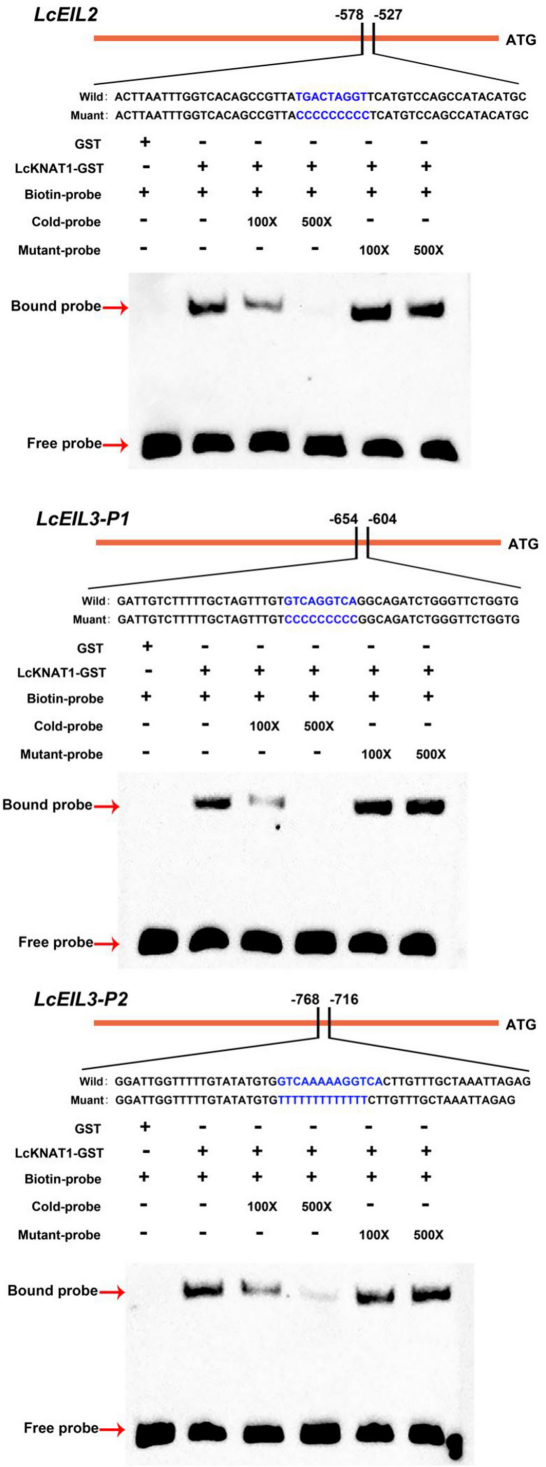


FIGURE 3 | Binding of LcKNAT1 to *LcEIL2* and *LcEIL3* promoters *in vitro*. The electrophoretic mobility shift assay (EMSA) using purified LcKNAT1-GST fusion protein. The sequences of wild and mutant probes were shown at the top of the images. Shifted bands, suggesting the formation of DNA-protein complexes, are indicated by arrows. " – " represents absence, and " + " represents presence. 100 × and 500 × indicate the increasing concentrates of probes. Non-biotin-labeled probe with the same sequence was used as a cold competitor. Only the GST protein was served as a negative control.

after two fruitlet abscission-induced approaches, namely, girdling plus defoliation and ET treatments (Ma et al., 2020). To obtain more detailed information regarding their expression patterns, we performed fruitlet removal treatment in this study, and their expression was monitored starting from 1 h after treatment. As shown in **Figure 1A**, peduncle abscission was induced 48 h after treatment, possibly due to the continuous polar flow of auxin passing through the peduncle AZ was eliminated by the removal of the fruitlet (Meir et al., 2010; Kućko et al., 2020). At 96 h after fruitlet removal, more than 89% of peduncles were abscised. Consistent with our previous results, *LcKNAT1* was sharply reduced from 1 h after fruitlet removal and then maintained this low expression level throughout the abscission process (**Figure 1B**). In contrast, both *LcEIL2* and *LcEIL3* were significantly increased from 6 h after fruitlet removal (**Figures 1C,D**). Together, these results indicate that the expression pattern of *LcEIL2/3* is opposite to that of *LcKNAT1* during the fruitlet abscission process in litchi.

LcKNAT1 Binds to the Promoters of *LcEIL2/3* and Represses Their Expression

Through screening the *cis*-elements in the promoter sequences of *LcEIL2/3*, we found that both *LcEIL2* and *LcEIL3* promoters contained the TGAC motifs, a KNOX TF-binding motif (Bolduc et al., 2012; **Figure 2A**). Thus, these findings prompted us to examine whether *LcEIL2/3* promoters could be recognized by LcKNAT1. To test it, we first performed a Y1H assay using the LcEIL2/3 promoters as baits and LcKNAT1 as prey. As shown in **Figure 2B**, the yeast cells containing the promoter of *LcEIL2* or *LcEIL3* failed to grow on SD/-Leu medium with AbA. In contrast, the yeast cells co-transformed with LcKNAT1, and the promoters of *LcEIL2* or *LcEIL3* could grow well on SD/-Leu medium with AbA, suggesting that LcKNAT1 can bind to the promoters of *LcEIL2* or *LcEIL3* and activate the reporter genes in yeast. We then carried out electrophoretic mobility shift assays (EMSAs) using GST-LcKNAT1 fusion protein. DNA fragments (∼50 bp) containing KNOX *cis*-element in the promoter regions of *LcEIL2/3* were synthesized and labeled with biotin at the 3′ end. As shown in **Figure 3**, when GST-LcKNAT1 fusion protein was incubated with labeled probes, the recombinant LcKNAT1 was able to strongly bind to the *LcEIL2* or *LcEIL3* promoter fragments, and this binding could be abolished by high concentrates of unlabeled competitors containing the same sequences in a dosage-dependent manner but not by competitors containing the mutant binding sites. These findings indicate that LcKNAT1 directly and specifically binds to the TGAC element within the promoter of *LcEIL2/3 in vitro*.

Furthermore, we performed a dual-LUC assay to examine whether LcKNAT1 could repress the activity of *LcEIL2/3* promoters. As shown in **Figure 4**, compared with control,

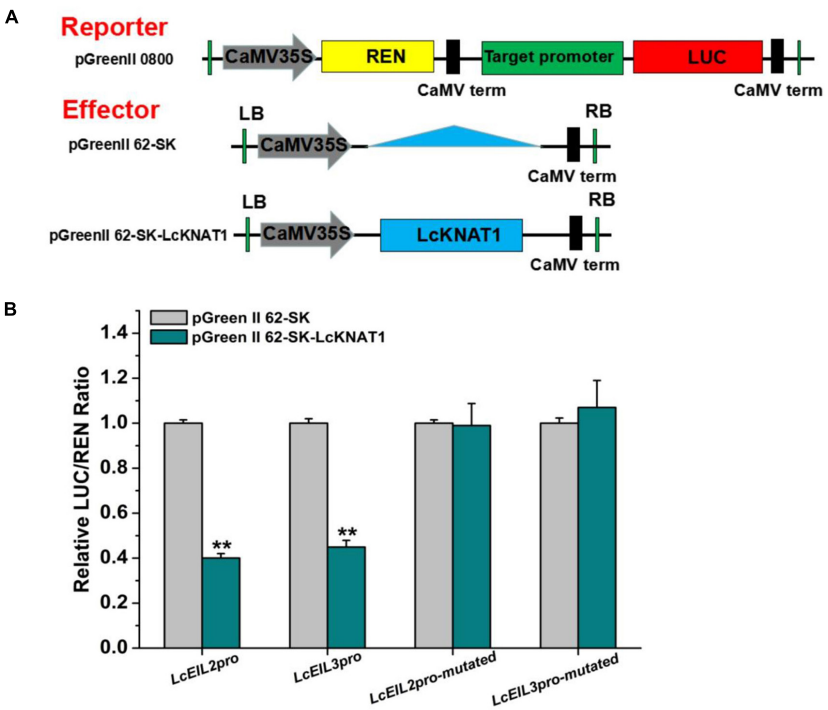


FIGURE 4 | LcKNAT1 represses the transcriptional activity of *LcEIL2* and *LcEIL3*. **(A)** Schematic representation of the double reporter and effector plasmids used for LcKNAT1 transcriptional activity assay on target gene promoters. **(B)** Transient expression assays on LcKNAT1 transcriptional repression of *LcEIL2* and *LcEIL3*. The reporter and effector vectors were co-introduced into tobacco leaves. After 48-h incubation, the suppression of *LcEIL2/3* and mutated *LcEIL2/3* promoters by LcKNAT1 was shown by the ratio of LUC/REN. Mutated *LcEIL2/3* promoters were as shown in **Figure 3**. Each value represents the means of six biological replicates, and vertical bars represent ± SD. Asterisks indicate a significant difference (Student's *t*-test: $^{**}P < 0.01$).

the LUC/REN ratio was dramatically reduced when either the Pro_{LcEIL2}-LUC or the Pro_{LcEIL3}-LUC reporter construct was co-transfected with the effector of pGreenII 62-SK-LcKNAT1, indicating that LcKNAT1 could repress the promoter activity of *LcEIL2/3*. Additionally, this repression was abolished when the core TGAC elements of *LcEIL2/3* promoters were mutated, indicating that LcKNAT1 can bind to the promoters of *LcEIL2/3 via* the core TGAC element and repress their activity.

The Ectopic Expression of *LcKNAT1* in *Arabidopsis* Delays the Floral Organ Abscission and Suppresses the Promoter Activity of *LcEIL2/3*

To further confirm that *LcEIL2/3* is negatively regulated by LcKNAT1, we carried out a stable transformation assay in *Arabidopsis*. First, we generated three transgenic *Arabidopsis* lines expressing *LcKNAT1* (**Figure 5A**). *35S:LcKNAT1-1* and *35S:LcKNAT1-2* transgenic lines showed a higher expression level and were selected for further functional analysis. As shown in **Figure 5B**, for wild-type Col, when plants bore 30 siliques, there were only two siliques with floral organs attached. In contrast, *35S:LcKNAT1-1* and *35S:LcKNAT1-2* transgenic plants still retained the floral organs at position 30. The increase of cytosolic pH at AZ was reported to be closely associated with organ abscission in *Arabidopsis* (Sundaresan et al., 2014). We then used a pH-sensitive indicator, BCECF-AM, to detect the pH value in the cytosol of AZ tissues of transgenic plants. We found that BCECF signals were obviously appeared at position 4 in Col, while the BCECF signals could not be detected in both *35S:LcKNAT1-1* and *35S:LcKNAT1-2* transgenic lines, consistent with their floral organ abscission process (**Figure 5C**). Taken

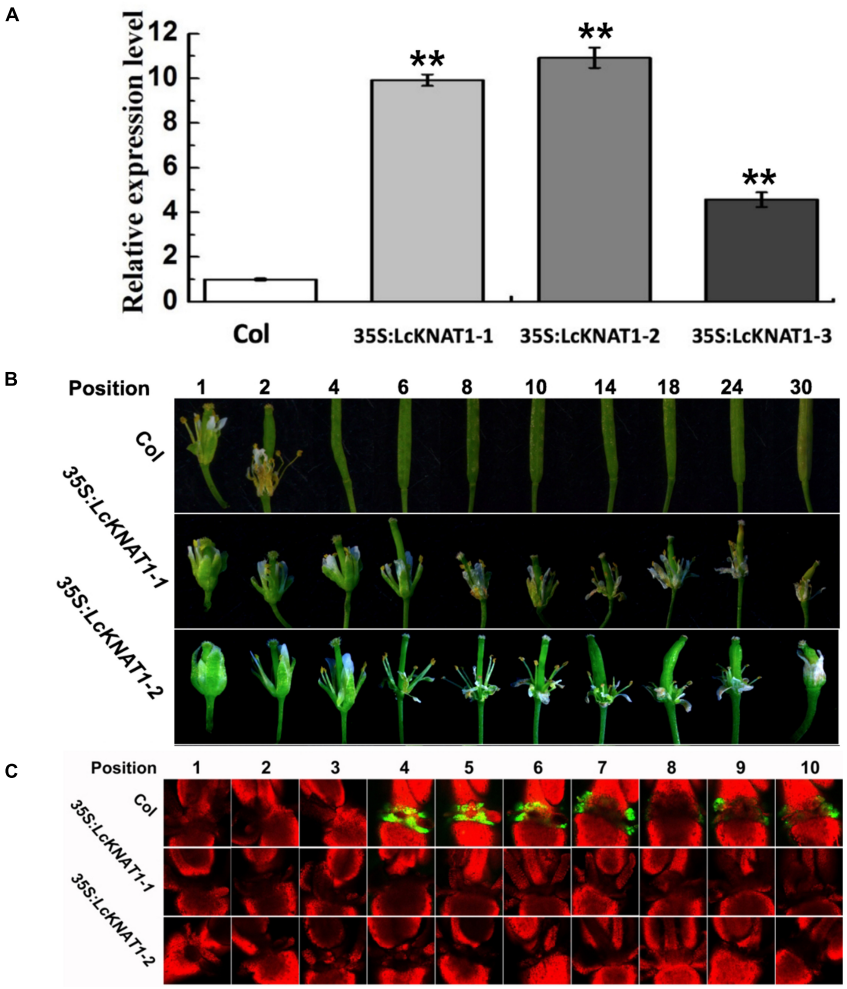


FIGURE 5 | Ectopic expression of *LcKNAT1* in *Arabidopsis* suppresses floral organ abscission. **(A)** The expression levels of *LcKNAT1* in *Arabidopsis* transgenic lines. *LcKNAT1* driven by the CaMV 35S promoter was transformed into wild-type *Arabidopsis*. **(B)** Phenotypes of floral organ abscission process in Col and *LcKNAT1* transgenic lines. Floral organ position numbers were counted from the first flower with visible white petals on the top of the inflorescence. **(C)** BCECF fluorescence micrographs of floral organ AZ in Col and *LcKNAT1* transgenic lines. Inflorescence samples were separated from the plant body, incubated in BCECF-AM staining, and then examined using the confocal laser scanning microscope. Merged images of BCECF fluorescence with chlorophyll autofluorescence were shown, which are representative images out of 3–4 replicates. Asterisks indicate a significant difference (Student's *t*-test: $^{**}P < 0.01$).

together, these findings suggest that LcKNAT1 functions as a strong negative regulator in control of the floral organ abscission in *Arabidopsis*.

Then, we ectopically expressed the GUS reporter driven by the promoter of *LcEIL2* or *LcEIL3* in *Arabidopsis*. GUS staining revealed that both *LcEIL2* and *LcEIL3* had a dominant accumulation at the floral AZ. A stronger GUS signal in floral AZ could be detected from positions 3 to 9 in the Pro_{LcEIL2}*:GUS* plant and positions 5 to 9 in Pro_{LcEIL3}*:GUS* plant, respectively. Then, homozygous Pro_{LcEIL2}*:GUS* plants or Pro_{LcEIL3}*:GUS* plants were crossed with homozygous *35S:LcKNAT1-2* plants, and F1 *35S:LcKNAT1-2* × Pro_{LcEIL2}*:GUS* plants or *35S:LcKNAT1-2* × Pro_{LcEIL3}*:GUS* plants were used for GUS staining. As shown in **Figure 6A**, both the GUS signals driven by the promoter of *LcEIL2* or *LcEIL3* at the positions detected were significantly suppressed in the presence of *LcKNAT1*. Additionally, relative quantitative assays show that the expression level of the *GUS* gene driven by *LcEIL2* or *LcEIL3* promoter was obviously downregulated by *LcKNAT1*, with 7.3-fold and 6.9-fold reduction, respectively (**Figure 6B**), further supporting that LcKNAT1 could suppress the promoter activity of *LcEIL2*/*3 in planta*.

DISCUSSION

It is well known that ET is a strong inducer of organ abscission and has been widely reported in many fruit crops (Nunezelisea, 1986; Rasori et al., 2002; Dal et al., 2005; Gil-Amado and Gomez-Jimenez, 2012; Uzquiza et al., 2014; Li et al., 2015; Merelo et al., 2017). However, the mechanisms underlying ET in the regulation of abscission are largely unknown. In a previous study, we revealed that LcKNAT1 acts as a negative regulator of fruitlet abscission by repressing ET

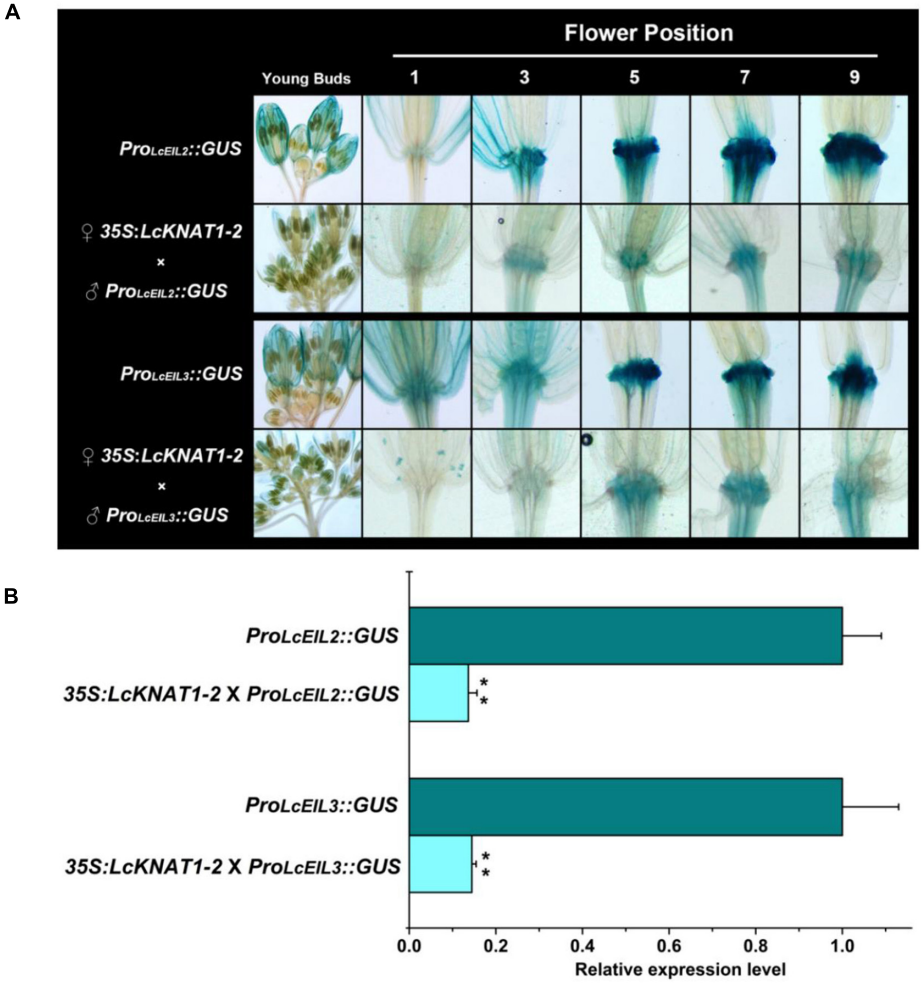


FIGURE 6 | GUS expression driven by the promoter of *LcEIL2* or *LcEIL3* is repressed by LcKNAT1. **(A)** ♂Pro_{LcEIL2}*:GUS* × ♀*35S:LcKNAT1-2* indicates that T3 homozygous Pro_{LcEIL2}*:GUS* plant was crossed with T3 homozygous *35S:LcKNAT1-2* plant, and F1 *35S:LcKNAT1-2* × Pro_{LcEIL2}*:GUS* or *35S:LcKNAT1-2* × Pro_{LcEIL3}*:GUS* plants were used for GUS staining. **(B)** Quantitative assay of the GUS expression level in Pro_{LcEIL2}*:GUS* plants, *35S:LcKNAT1-2* × Pro_{LcEIL2}*:GUS* plants, Pro_{LcEIL3}*:GUS* plants, and *35S:LcKNAT1-2* × Pro_{LcEIL3}*:GUS* plants, respectively. Asterisks indicate a significant difference (Student's *t*-test: $^{**}P < 0.01$).

biosynthetic genes (Zhao et al., 2020). In this study, we further proved that the ET signaling components *LcEIL2/3* were also repressed directly by LcKNAT1 using both transient expression systems and stable transformation assays. We thus propose that a LcKNAT1-LcEIL2/3 regulatory module might be involved in the fruit abscission in litchi.

Results from the ectopic expression of *LcKNAT1* in tomatoes and *Arabidopsis* implied that LcKNAT1 likely played critical roles in the control of the fruitlet abscission in litchi. In combination with the findings from our previous and this study, how LcKNAT1 functions as a strong repressor during the fruitlet abscission in litchi could be partially explained as follows: (i) ET is the key inducer of abscission, and *LcACS1/7* and *LcACO2/3* involved in ET synthesis are induced at the fruitlet AZ; (ii) LcKNAT1 can directly repress the ET biosynthetic genes including *LcACS1/7* and *LcACO2/3*, to suppress ET production; (iii) LcKNAT1 can directly repress the transcription of *LcEIL2/3*, which are the central regulators of ET signaling; (iv) LcEIL2/3 can directly activate the transcription of *LcACS1/4/7* and *LcACO2/3* and cell wall degradation genes *LcCEL2/8* and *LcPG1/2*. Additionally, ET can repress *LcKNAT1* expression but promote *LcEIL2/3* transcription. Therefore, before abscission is triggered, *LcKNAT1* expression is maintained at a relatively high level, thereby resulting in a repressive state of both ET biosynthesis and signaling. In contrast, when *LcKNAT1* expression is inhibited by abscission-inducing signals, both the ET biosynthesis and signaling are derepressed, thereby activating the abscission process (**Figure 7**).

In the model plant *Arabidopsis*, genetic evidence has demonstrated that BP/KNAT1 functions in the control of the floral organ abscission by regulating the activity of two other KNOX-related genes *KNAT2* and *KNAT6* (Shi et al., 2011). In our previous study, a total of eight KNOX homologs were identified from the litchi genome. Interestingly, only LcKNAT1 displayed a regulatory function in the control of the floral organ abscission in *Arabidopsis* (Zhao et al., 2020). However, the overexpression of *KNAT2* or *KNAT6* only slightly promoted the floral organ abscission in *Arabidopsis* wild-type background (Shi et al., 2011), and it is possible that the overexpression of *LcKNAT3* or *LcKNAT6*, as LcKNAT3 and LcKNAT6 were the homologs to AtKNAT2 and AtKNAT6, displayed very weak phenotype. It will be interesting to further check whether the overexpression of *LcKNAT3* or *LcKNAT6* could also rescue the abscission phenotype in *ida* mutant. Thus, so far it could not be claimed that the mechanism underlying LcKNAT1 in control of the fruit abscission in litchi is different from that of BP/KNAT1 in *Arabidopsis*. Furthermore, we showed previously that KNAT1 binding motifs were also conserved in *Arabidopsis*; however, role of KNAT1 in regulating ET synthetic genes has never been investigated. It is likely that KNAT1 could negatively regulate both other KNOX members and the genes related to ET biosynthesis and signaling during the fruit abscission in litchi or floral organ abscission in *Arabidopsis*.

In *Arabidopsis*, BP/KNAT1 acts as the key downstream TF of IDA-HAE/HSL2 signaling to repress the floral organ abscission (Cho et al., 2008; Stenvik et al., 2008; Shi et al., 2011). Previously, we also found that the homologs of IDA and HAE/HSL2 are present in the litchi genome and demonstrated that LcIDL1 and LcHSL2 have similar functions in the control of the floral organ abscission in *Arabidopsis* (Ying et al., 2016; Wang et al., 2019). Additionally, since the ectopic expression of *LcKNAT1* decreases pedicel AZ activation in tomatoes (Zhao et al., 2020) and blocks floral organ abscission in *Arabidopsis* (**Figure 5B**), we can speculate that a conserved LcIDL1-LcHSL2 signaling module is probably sitting upstream and regulating the repressor activity of LcKNAT1. Interestingly, *ida* mutants could not be rescued by ET treatment, which led to the conclusion that the IDA-HAE/HSL2 pathway may control abscission in an ET-independent manner

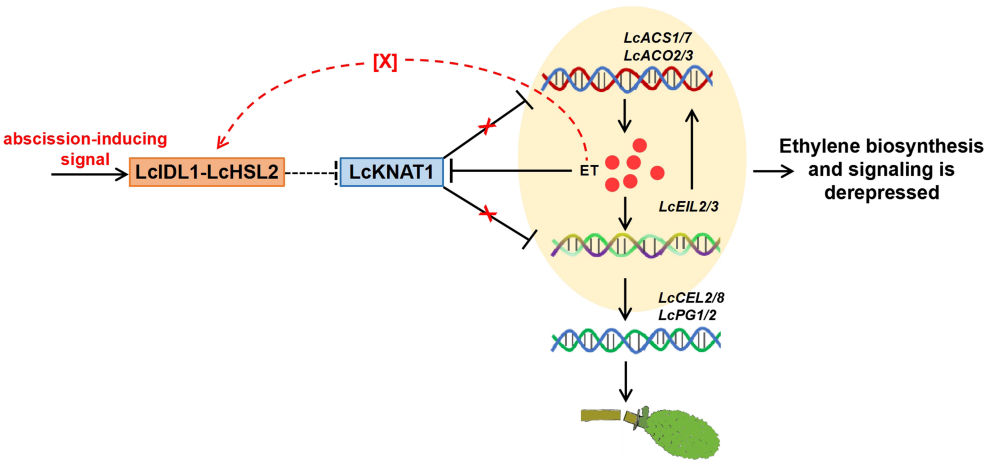


FIGURE 7 | Proposed model for the role of LcKNAT1 in ethylene (ET)-regulated fruitlet abscission in litchi. LcKNAT1 can directly repress *LcACS1/7*, *LcACO2/3*, and *LcEIL2/3*, when LcKNAT1 is repressed by unknown abscission-inducing signals *LcACS1/7* and *LcEIL2/3* which are derepressed. Furthermore, LcEIL2/3 can also bind to *LcACS1/7* and *LcACO2/3* but promote their transcription, therefore forming a feedback loop to enhance the ET biosynthesis and signaling. Additionally, ET could also disable the repressor function of LcKNAT1 through the indirect activation of the LcIDL1-LcHSL2 module *via* an unknown factor, thereby enhancing the activation of the cell wall degradation genes *LcCEL2/8* and *LcPG1/2* to induce fruitlet abscission.

(Jinn et al., 2000; Butenko et al., 2003, 2006). However, this conclusion has been recently reconsidered with the findings that the IDA homologs from different species, including tomato, soybean, oil palm, citrus, and yellow lupine, are induced during the abscission process activated by ET (Tucker and Yang, 2012; Estornell et al., 2015; Stø et al., 2015; Wilmowicz et al., 2018). In litchi, *LcIDL1* and *LcHSL2* in the AZ are also ET-induced during the fruit abscission (Ying et al., 2016; Wang et al., 2019). Our results partially support the conclusion that the IDA-HAE/HSL2 abscission signaling module is ET-dependent (Botton and Ruperti, 2019; Meir et al., 2019) and propose that LcKNAT1 may act as the downstream of LcIDL1-LcHSL2 signaling subject to a no direct positive ET feedback mechanism and is required for the last steps of the execution of the abscission process. Thus, we hypothesized that ET treatment could have a dual effect by triggering the expression of genes encoding cell wall hydrolases and disabling the repressor function of LcKNAT1 through the indirect activation of the LcIDL1-LcHSL2 module *via* an unknown factor (**Figure 7**). In the future, efforts should be devoted to explore the role of ET in the AZ, in particular the cross talk between ET and IDA-HAE/HSL2 signaling pathways.

DATA AVAILABILITY STATEMENT

The authors acknowledge that the data presented in this study must be deposited and made publicly available in an acceptable repository, prior to publication. Frontiers cannot accept an article that does not adhere to our open data policies.

AUTHOR CONTRIBUTIONS

MZ and JL conceived and designed the experiments. XM performed most of the experiments. PY, ZH, and HW provided assistance. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the National Natural Science Foundation of China (32072514 and 32072544), the Laboratory of Lingnan Modern Agriculture Project (NZ NT2021004), and the National Key R&D Program of China (2020YFD1000101).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.802016/full#supplementary-material

REFERENCES

Bolduc, N., Yilmaz, A., Mejia-Guerra, M. K., Morohashi, K., O'Connor, D., Grotewold, E., et al. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. *Genes Dev.* 26, 1685–1690. doi: 10.1101/gad.193433.112

Botton, A., and Ruperti, B. (2019). The yes and no of the ethylene involvement in abscission. *Plants* 8:187. doi: 10.3390/plants8060187

Butenko, M. A., Patterson, S. E., Grini, P. E., Stenvik, G. E., Amundsen, S. S., Mandal, A., et al. (2003). Inflorescence deficient in abscission controls floral organ abscission in *Arabidopsis* and identifies a novel family of putative ligands in plants. *Plant Cell* 15, 2296–2307. doi: 10.1105/tpc.014365

Butenko, M. A., Stenvik, G. E., Alm, V., Saether, B., Patterson, S. E., and Aalen, R. B. (2006). Ethylene-dependent and-independent pathways controlling floral abscission are revealed to converge using promoter::reporter geneconstructs in the *ida* abscission mutant. *J. Exp. Bot.* 57, 3627–3637. doi: 10.1093/jxb/erl130

Cho, S. K., Larue, C. T., Chevalier, D., Wang, H., Jinn, T. L., Zhang, S., et al. (2008). Regulation of floral organ abscission in *Arabidopsis* thaliana. *Proc. Nat. Acad. Sci. USA* 105, 15629–15634. doi: 10.1073/pnas.0805539105

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium- mediated transformation of *Arabidopsis* thaliana. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Dal, C., Danesin, M., Andrea, B., Dorigoni, A., and Ramina, A. (2005). Ethylene biosynthesis and perception in apple fruitlet abscission Malus domestica L. Borck Valeriano. *J. Exp. Bot.* 56, 2995–3005. doi: 10.1093/jxb/eri296

Estornell, L. H., Agusti, J., Merelo, P., Talon, M., and Tadeo, F. R. (2013). Elucidating mechanisms underlying organ abscission. *Plant Sci.* 199, 48–60. doi: 10.1016/j.plantsci.2012.10.008

Estornell, L. H., Wildhagen, M., Pérez-Amador, M. A., Talón, M., Tadeo, F. R., and Butenko, M. A. (2015). The IDA peptide controls abscission in *Arabidopsis* and *Citrus*. *Front. Plant Sci.* 6:1003. doi: 10.3389/fpls.2015.01003

Gil-Amado, J. A., and Gomez-Jimenez, M. C. (2012). Regulation of polyamine metabolism and biosynthetic gene expression during olive mature-fruit abscission. *Planta* 235, 1221–1237. doi: 10.1007/s00425-011-1570-1

Gomez-Cadenas, A., Tadeo, F. R., Talon, M., and Primo-Millo, E. (1996). Leaf abscission induced by ethylene in water-stressed intact seedlings of cleopatra mandarin requires previous abscisic acid accumulation in roots. *Plant Physiol.* 112, 401–408. doi: 10.1104/pp.112.1.401

Jinn, T. L., Stone, J. M., and Walker, J. C. (2000). HAESA, an *Arabidopsis* leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes Dev.* 14, 108–117.

Kitsaki, C. K., Drossopoulos, J. B., Aivalakis, G., Anastasiadou, F., and Delis, C. (1999). *In vitro* studies of ABA and ethephon induced abscission in olive organs. *J. Pomol. Horticult. Sci.* 74, 19–25.

Kućko, A., Wilmowicz, E., Pokora, W., and Alché, J. (2020). Disruption of auxin gradient in abscission zone area evokes asymmetrical changes leading to flower separation in yellow lupine. *Int. J. Mol. Sci.* 21:3815. doi: 10.3390/ijms21113815

Li, C., Wang, Y., Ying, P., Ma, W., and Li, J. (2015). Genome-wide digital transcript analysis of putative fruitlet abscission related genes regulated by ethephon in litchi. *Front. Plant Sci.* 6:502. doi: 10.3389/fpls.2015.00502

Ma, X. S., Li, C. Q., Huang, X. M., Wang, H. C., Wu, H., Zhao, M. L., et al. (2019). Involvement of HD-ZIP I transcription factors LcHB2 and LcHB3 in fruitlet abscission by promoting transcription of genes related to the biosynthesis of ethylene and ABA in litchi. *Tree Physiol.* 39, 1600–1613.

Ma, X. S., Yuan, Y., Li, C. Q., Wu, Q., He, Z. D., Li, J. G., et al. (2021). Brassinosteroids supress ethylene-induced fruitlet abscission through LcBZR1/2-mediated transcriptional repression of *LcACS1/4* and *LcACO2/3* in litchi. *Horticult. Res.* 8:105. doi: 10.1038/s41438-021-00540-z

Ma, X. S., Yuan, Y., Wu, Q., Wang, J., Li, J. G., and Zhao, M. L. (2020). LcEIL2/3 are involved in fruitlet abscission via activating genes related to ethylene biosynthesis and cell wall remodeling in litchi. *Plant J.* 103, 1338–1350. doi: 10.1111/tpj.14804

Meir, S., Philosoph-Hadas, S., Rivo, J., Tucker, M. L., Patterson, S. E., and Roberts, J. A. (2019). Re-evaluation of the ethylene-dependent and -independent pathways in the regulation of floral and organ abscission. *J. Exp. Bot.* 70, 1461–1467. doi: 10.1093/jxb/erz038

Meir, S., Philosoph-Hadas, S., Sundaresan, S., Selvaraj, K. S., Burd, S., Ophir, R., et al. (2010). Microarray analysis of the abscission- related transcriptome in the

tomato flower abscission zone in response to auxin depletion. *Plant Physiol.* 154, 1929–1956. doi: 10.1104/pp.110.160697

Meng, J., Zhou, Q., Zhou, X., Fang, H. X., and Ji, S. J. (2019). Ethylene and 1-MCP treatments affect leaf abscission and associated metabolism of Chinese cabbage. *Postharvest Biol. Technol* 157:963.

Merelo, P., Agusti, J., Arbona, V., Costa, M. L., Estornell, L. H., Gomez-Cadenas, A., et al. (2017). Cell wall remodeling in abscission zone cells during ethylene-promoted fruit abscission in citrus. *Front. Plant Sci.* 8:126.

Mitra, S. K., Pereira, L. S., Pathak, P. K., and Majumdar, D. (2003). *Fruit abscission pattern of lychee cultivars. In 2nd International Symposium on Lychee, Longan, Rambutan and other Sapindaceae Plants (Chiang Mai)*. Leuven: ISHS. 215–218.

Nunez-Elisea, R., and Davenport, T. L. (1986). Abscission of mango fruitlets as influenced by enhanced ethylene biosynthesis. *Plant Physiol.* 82, 991–994. doi: 10.1104/pp.82.4.991

Nunezelisea, R. (1986). Abscission of mango fruitlets as influenced by enhanced ethylene biosynthesis. *Plant Physiol.* 82, 991–994.

Parra-Lobato, M. C., and Gomez-Jimenez, M. C. (2011). Polyamine-induced modulation of genes involved in ethylene biosynthesis and signaling pathways and nitric oxid production during olive mature fruit abscission. *J. Exp. Bot.* 62, 4447–4465. doi: 10.1093/jxb/err124

Patterson, S. E. (2001). Cutting loose Abscission and dehiscence in *Arabidopsis*. *Plant Physiol.* 126, 494–500. doi: 10.1104/pp.126.2.494

Rasori, A., Ruperti, B., Bonghi, C., Tonutti, P., and Ramina, A. (2002). Characterization of two putative ethylene receptor genes expressed during peach fruit development and abscission. *J. Exp. Bot.* 53, 2333–2339. doi: 10.1093/jxb/erf097

Roberts, J. A., Elliott, K. A., and Gonzalez-Carranza, Z. H. (2002). Abscission, dehiscence, and other cell separation processes. *Ann. Rev. Plant Biol.* 53, 131–158. doi: 10.1146/annurev.arplant.53.092701.180236

Santiago, J., Brandt, B., Wildhagen, M., Hohmann, U., Hothorn, L. A., Butenko, M. A., et al. (2016). Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. *Elife* 5:e15075. doi: 10.7554/eLife.15075

Sexton, R., and Roberts, J. A. (1982). Cell biology of abscission. *Ann. Rev. Plant Physiol.* 33, 133–162.

Shi, C. L., Stenvik, G. E., Vie, A. K., Bones, A. M., Pautot, V., Proveniers, M., et al. (2011). *Arabidopsis* class I KNOTTED-like homeobox proteins act downstream in the IDA- HAE/HSL2 floral abscission signaling pathway. *Plant Cell* 23, 2553–2567. doi: 10.1105/tpc.111.084608

Stenvik, G. E., Tandstad, N. M., Guo, Y., Shi, C. L., Kristiansen, W., Holmgren, A., et al. (2008). The EPIP peptide of INFLORESCENCE DEFICIENT IN ABSCISSION is sufficient to induce abscission in *Arabidopsis* through the receptor-like kinases HAESA and HAESA-LIKE2. *Plant Cell* 20, 1805–1817. doi: 10.1105/tpc.108.059139

Stø, I. M., Orr, R. J. S., Fooyontphanich, K., Jin, X., Knutsen, J. M. B., Fischer, U., et al. (2015). Conservation of the abscission signaling peptide IDA during Angiosperm evolution: Withstanding genome duplications and gain and loss of the receptors HAE/HSL2. *Front. Plant Sci.* 6:931. doi: 10.3389/fpls.2015.00931

Sundaresan, S., Philosoph-Hadas, S., Riov, J., Belausov, E., Kochanek, B., Tucker, M. L., et al. (2014). Abscission of flowers and floral organs is closely associated with alkalization of the cytosol in abscission zone cells. *J. Exp. Bot.* 66, 913–919. doi: 10.1093/jxb/eru483

Taylor, J. E., and Whitelaw, C. A. (2001). Signals in abscission. *New Phytol.* 151, 323–340.

Tieman, D. M., Ciardi, J. A., Taylor, M. G., and Klee, H. J. (2001). Members of the tomato *LeEIL* (EIN3-like) gene family are functionally redundant and regulate ethylene responses throughout plant development. *Plant J.* 26, 47–58. doi: 10.1046/j.1365-313x.2001.01006.x

Tucker, M. L., and Yang, R. (2012). IDA-like gene expression in soybean and tomato leaf abscission and requirement for a diffusible stelar abscission signal. *AoB Plants* 2012:ls035. doi: 10.1093/aobpla/pls035

Uzquiza, L., Martin, P., Sievert, J. R., Arpaia, M. L., and Fidelibus, M. W. (2014). Methyl jasmonate and 1-aminocyclopropane-1-carboxylic acid interact to promote grape berry abscission. *Am. J. Enol. Viticult.* 65, 504–509.

Wang, F., Zheng, Z. H., Yuan, Y., Li, J. G., and Zhao, M. L. (2019). Identification and characterization of HAESA-Like genes involved in the fruitlet abscission in Litchi. *Int. J. Mol. Sci.* 20:5945. doi: 10.3390/ijms20235945

Wilmowicz, E., Frankowski, K., Kućko, A., Świdziński, M., de Dios Alché, J., Nowakowska, A., et al. (2016). The influence of abscisic acid on the ethylene biosynthesis pathway in the functioning of the flower abscission zone in *Lupinus luteus*. *J. Plant Physiol.* 206, 49–58. doi: 10.1016/j.jplph.2016.08.018

Wilmowicz, E., Kućko, A., Ostrowski, M., and Panek, K. (2018). INFLORESCENCE DEFICIENT IN ABSCISSION-like is an abscission associated and phytohormone-regulated gene in flower separation of *Lupinus luteus*. *Plant Growth Regul.* 85, 91–100.

Yang, S. F., and Hoffman, N. E. (1984). Ethylene biosynthesis and its regulation in higher plants. *Ann. Rev. Plant Physiol.* 35, 155–189.

Ying, P., Li, C., Liu, X., Xia, R., Zhao, M., and Li, J. (2016). Identification and molecular characterization of an IDA-like gene from litchi, *LcIDL1*, whose ectopic expression promotes floral organ abscission in *Arabidopsis*. *Scient. Rep.* 6:37135. doi: 10.1038/srep37135

Yuan, R., and Huang, H. B. (1988). Litchi fruit abscission: its patterns, effect of shading and relation to endogenous abscisic acid. *Sci. Horticult.* 36, 281–292.

Zhong, H. Y., Chen, J. W., Li, C. Q., Chen, L., Wu, J. Y., Chen, J. Y., et al. (2011). Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions. *Plant Cell Rep.* 30, 641–653. doi: 10.1007/s00299-010-0992-8

Zhao, M., Li, C., Ma, X., Xia, R., Chen, J., Liu, X., et al. (2020). KNOX protein KNAT1 regulates fruitlet abscission in litchi by repressing ehtylene biosynthetic genes. *J. Exp. Bot.* 71, 4069–4082. doi: 10.1093/jxb/eraa162

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fpls.2022.823993

Morphological and Developmental Features of Stone Cells in *Eriobotrya* Fruits

Shoukai Lin[1,2†], Dahe Lin[1,2†], Bisha Wu[1,2], Shiwei Ma[1,2], Shengfeng Sun[1], Ting Zhang[1], Wenting Zhang[1], Yunlu Bai[3], Qiong Wang[3] and Jincheng Wu[1,2]*

[1]College of Environmental and Biological Engineering, Putian University, Putian, China, [2]Fujian Provincial Key Laboratory of Ecology-Toxicological Effects & Control for Emerging Contaminants, Putian University, Putian, China, [3]College of Horticulture, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Hao Yu,
National University of Singapore, Singapore

Reviewed by:
Kui Lin-Wang,
The New Zealand Institute for Plant and Food Research Ltd., New Zealand
Yongping Cai,
Anhui Agricultural University, China

***Correspondence:**
Jincheng Wu
wjc2384@163.com

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

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

Received: *28 November 2021*
Accepted: *06 January 2022*
Published: *27 January 2022*

Citation:
Lin S, Lin D, Wu B, Ma S, Sun S, Zhang T, Zhang W, Bai Y, Wang Q and Wu J (2022) Morphological and Developmental Features of Stone Cells in Eriobotrya Fruits. Front. Plant Sci. 13:823993. doi: 10.3389/fpls.2022.823993

Some members of the *Rosaceae* family, particularly pear, contain stone cells in their fruits. Although stone cells in pear fruits are well studied, relatively little attention has been given to loquat stone cells. Only a few reports have suggested a relationship between stone cell traits and storage and transport tolerance of loquat fruits. Previously, we generated the variety JT8 from the interspecific hybrid of the loquat cultivar Jiefangzhong (JFZ; *Eriobotrya japonica* Lindl. cv. Jiefangzhong, female parent) and wild Taiwanese loquat (TL; *E. deflexa* Nakai, male parent). The JT8 fruits had a granular feel, similar to that of pear fruits, due to the presence of stone cells. In this study, the shape, size, development, and distribution dynamics of stone cells of *Eriobotrya* plants were thoroughly investigated. The results showed that loquat stone cells are brachysclereids and often contain typical branching pits. Loquat stone cells were distributed as both single stone cells and in stone cell clusters (SCCs), and the density of the stone cells near the core was higher than that near the peel. Stone cell density first increased and then decreased during fruit development. These traits noted in *Eriobotrya* were very similar to those observed in pear, indicating a close relationship between loquat and pear. Moreover, the contents, density dynamics, and aggregation traits of stone cells of the interspecific hybrid JT8 were derived from the male parent (TL). Transgressive segregation was likely exhibited in the content of stone cells and the size of the SCCs. More specifically, the content of stone cells reached 1.61% (w/w). In extreme cases, SCCs of JT8 exceeded 1,000 μm in length and 500 μm in width. This demonstrated that stone cell traits could be transmitted from parent to progeny through interspecific hybridization. The density dynamics of stone cells in two loquat cultivars with different storage and transport tolerances were also investigated, which indicated that the cultivar with more stone cells was more tolerant to storage and transport. We suggest that wild loquat genetic resources containing stone cells in *Eriobotrya* plants can be used to gradually improve the storage and transport tolerance of loquat fruits.

Keywords: *Eriobotrya* plants, stone cells, fruits, morphological and developmental features, trait transmission

INTRODUCTION

Loquat (*Eriobotrya japonica* Lindl.), a characteristic evergreen fruit tree species originating from China, is now distributed in more than 30 countries, mainly in East Asian, South Asian, and coastal Mediterranean countries. Loquat fruits are not easy to store and transport, and the period during which fresh fruits are available is short, which restricts the improvement of the planting efficiency of loquat (Lin et al., 2018). Some kinds of *Rosaceae* fruits contain numerous stone cells, such as pear and haw fruits, which make a distinct impression to many people. Although careful researchers have noted that some loquat cultivars also contain stone cells distributed within their fruits, little attention has been given to stone cells in loquat, which also belongs to *Rosaceae*. Stone cells were larger and more abundant in the storage- and transport-tolerant cultivars Jiefangzhong, Wugongbai, and Ruoyangqing and smaller and rarer in the storage- and transport-susceptible cultivars Changhong No. 3, Baili, and Baisha, suggesting that stone cells exist in ripe fruits of some loquat varieties and are related to the storage and transport tolerance of loquat fruits (Lin et al., 2009; Zhu et al., 2018).

Stone cells (or sclereids) are sclerenchyma cells formed by the secondary deposition of lignin on the primary cell wall of parenchyma cells. Their development is closely related to the synthesis, transfer, and deposition of lignin. Stone cells can be distributed as single cells or as aggregates known as stone cell clusters (SCCs; Nii et al., 2008). To the best of our knowledge, many *Rosaceae* fruits contain stone cells, including those of *Pyrus* spp., *Eriobotrya japonica*, *Chaenomeles sinensis*, *Crataegus pinnatifida*, *Prunus mume*, *Prunus armeniaca*, *Prunus salicina*, and so forth (Huang et al., 2005; Sun and Li, 2006; Lin et al., 2009; Pan, 2011; Wu et al., 2013). In particular, the fruits of pear cultivars contain a large number of stone cells that significantly affect the quality of their fruits, thereby attracting much attention and making these cultivars model plants for the study of fruit stone cells (Wu et al., 2013). The literature thus far has shown that most of the studies on fruits containing stone cells are few and unthorough, except those pertaining to pear fruits. For instance, when describing the overall quality of hawthorn fruits, it was only mentioned that there were few stone cells (Pan, 2011).

In pear, secondary cell wall thickening and lignin deposition are key steps in the developmental process from parenchyma cells to stone cells. First, lignin is deposited along the apical angles of the primary cell wall and then gradually diffuses, which leads to uneven thickening of the cell walls. With the thickening of the secondary cell wall, the cell contents gradually shrink to the cell center. Eventually, the cell contents disappear, and the entire cell is completely composed of secondary cell wall structures (Choi and Lee, 2013; Zhao et al., 2013). Pear stone cells are brachysclereids that are characterized by thickened secondary cell walls with severe lignification, often with branching pits (Jin et al., 2013). The cell wall components of pear stone cells contain large amounts of lignin, cellulose, and hemicellulose (xylans), while the parenchyma cells of pear fruits contain abundant pectin (Brahem et al., 2017). The distribution of stone cells in pear flesh is not uniform and changes significantly in different growth stages (Choi and Lee, 2013). During fruit development, the density of stone cells first increased and then decreased, and the density of stone cells in different pear varieties reached its peak at different times (Zhang et al., 2017). Nii et al. (2008), Choi and Lee (2013), and Zhang et al. (2017) reported dynamic changes in stone cell distribution during the developmental process of pear fruits. The density of stone cells is higher near the core than near the peel. In addition, the number and size of stone cell clusters near the core are often higher than those near the peel (Tao et al., 2009). Different species and cultivars have diverse stone cell traits, such as content, shape, and size (Tao et al., 2009). When fruits can be used for traditional Chinese medicine, stone cell morphology is often used as one of the bases for the identification of medicinal varieties, such as *Crataegus pinnatifida* and *Prunus mume* (Huang et al., 2005; Sun and Li, 2006). Cao et al. (2010) found that the content of stone cells in pear pulp differed among 304 cultivars and interspecific hybrids. Tian et al. (2011) found that the size of SCCs in pear pulp differed among 319 cultivars and interspecific hybrids. Different stone cell traits are related to the texture and quality of pear fruits (Cao et al., 2010; Tian et al., 2011). There is a positive correlation between the content of pear stone cells and the firmness, adhesiveness, and chewiness of pear pulps (Choi et al., 2007; Liu et al., 2020).

In contrast, the fruits of loquat cultivars were similar to pear cultivars in the shape of stone cells, which belong to brachysclereids (Lin et al., 2009; Zhu et al., 2018). A series of recent studies reported the composition and development of loquat stone cells. It was found that loquat stone cells contained lignin, cellulose, and hemicellulose, the cell corner (CC) and middle lamella (ML) deposited only lignin and pectin, and parenchyma cells contained almost no lignin (Zhu et al., 2018, 2019). During the development of loquat stone cells, lignin and cellulose gradually filled stone cells, while pectin mainly filled the CC and ML. Loquat stone cells contained abundant lignin functional groups of coniferaldehyde and coniferyl alcohol (Huang et al., 2019). Stone cells and vascular bundles were the active areas of lignin deposition. Cyclic stone cells deposited lignin in both the inner and outer layers. When stone cells are fully filled with lignin, the outer layer of stone cells can still continue to deposit lignin so that adjacent parenchyma cells can also start to deposit lignin (Zhu et al., 2021). Therefore, some stone cells can be observed alone and surrounded by parenchyma cells. On the other hand, some parenchyma cells around stone cells also accumulate lignin, which later becomes stone cells and eventually forms stone cell clusters (Huang et al., 2019; Zhu et al., 2021). Interestingly, the newly synthesized lignin in loquat pulp was specifically deposited in the CC and ML of parenchyma cells around vascular bundles during storage, thus forming a network structure (Zhu et al., 2021).

The above studies indicated that pear stone cell traits were closely related to the texture and quality of pear fruits, which

has become an important index in pear breeding. In contrast, stone cell traits of loquat fruits are not well known, and some reports involving loquat stone cells suggested the existence of a certain relationship between stone cell traits and storage and transport tolerance in loquat fruits. Our former research work produced the hybrid variety JT8 from the interspecific hybrid of the common yellow-fleshed loquat cultivar Jiefangzhong (JFZ; *Eriobotrya japonica* Lindl. cv. Jiefangzhong, female parent) and wild Taiwanese loquat (TL; *E. deflexa* Nakai, male parent). It was found that the fruits of JT8 obviously had a stone cell taste similar to that of pear fruits. Although stone cell traits have potential application value to improve the storage and transport tolerance of loquat cultivars, the key problem of stone cell trait transmission from parents to progenies in interspecific crosses has not been solved in loquat. Therefore, in this study, the shape, size, development, and distribution dynamics of stone cells of *Eriobotrya* plants were thoroughly studied, and the potential breeding value of stone cell traits in improving the storage and transport tolerance of loquat fruits and the transmission of stone cell traits from parents to progeny in interspecific crosses were discussed.

MATERIALS AND METHODS

Plant Materials and Treatments

The fruits of *Eriobotrya* plants were used as plant materials, including the common loquat cultivars Jiefangzhong (JFZ; *Eriobotrya japonica* cv. Jiefangzhong) and Baili (BL; *E. japonica* cv. Baili), the wild Taiwanese loquat (TL; *E. deflexa* Nakai) and another form of Taiwanese loquat (TLk; *E. deflexa* f. *koshunensis* Nakai), and the following interspecific hybrids: JFZ (♀) × TL (♂) No. 8 (JT8) and JFZ (♀) × TLk (♂) No. 2 (JTk2) and No. 3 (JTk3). The fruits of cultivars JFZ and BL were derived from mature trees located in an outdoor orchard at 25°45' N and 118°55' E (Changtai town, Putian city, Fujian province, China). The fruits of other loquat plants were sampled from mature trees located in Loquat Plants Resource Nursery at 23°16' N and 113°37' E (South China Agricultural University, Guangzhou city, Guangdong province, China). The cultivars JFZ and BL underwent flower thinning at full bloom. Each inflorescence contained 10 ~ 15 fertilized ovaries with no petals and withered flower cores. Wild loquats and interspecific hybrids had few flower inflorescences; therefore, flower thinning was not carried out for these plants. The start of full bloom was defined as 0 DAF (days after full bloom). Three groups of samples were collected separately: the first group was used to determine the stone cell content and consisted of the mature fruits of JFZ, TL, TLk, JT8, JTk2, and JTk3; the second group was used to thoroughly investigate the shape, size, development, and distribution dynamics of stone cells and included the fruits of JFZ and JT8, which were sampled at 14, 28, 42, 56, 70, 84, 98, 112, and 126 DAF, and the fruit of TL, which was sampled on only four occasions at 14, 98, 112, and 126 DAF due to rare occurrence (in some years, no fruit was produced); and the third group was used to study the relationship between storability, fruit hardness, and stone cell development and included fruits of the storage- and transport-tolerant cultivar JFZ and susceptible cultivar BL (Lin et al., 2008), which were sampled at 14, 28, 63, 98, and 126 DAF.

Determination of the Stone Cell Content

The stone cell content of mature fruits was determined according to Nie's method with some modifications (Nie et al., 2006). The ripe loquat pulp was cut up and frozen in a freezer (−20°C) for 24 h. After thawing, the pulps were homogenized at 22,000 rpm for 3 min. The homogenized pulps were washed and incubated with 0.8 L distilled water for 3 min, and then, the supernatants were collected. The procedure was repeated several times until there was no sediment. The collected supernatants were passed through filter paper, and the stone cells remained on the filter paper. The collected stone cells were dried and weighed, and the content was calculated as follows: stone cell content (%) = weight of collected stone cells (g DW)/weight of fresh pulp (g FW) × 100.

Observation of Fruit Cross-Sections With Phloroglucinol-HCl Staining

Fresh loquat fruit was cut along the equatorial plane, and the cross-section was covered with 1 M HCl and incubated for 10 min. After slight drying with absorbent paper, the cross-section was covered with 1% phloroglucinol. After 10 min, the cross-section was observed by a SX-3 stereoscopic microscope (Shanghai Optical Instrument Factory, China) with a 20× microscopic field.

Observation of Frozen Sections With Phloroglucinol-HCl Staining

The fruit pulp was divided into small pieces of 0.5 mm × 0.5 mm × 0.5 mm and fixed with 4% paraformaldehyde (PFA) overnight. After that, the small pieces were embedded by using O.C.T. compound (SAKURA, Japan). Frozen tissue sections with a thickness of 30 μm were cut by a CM1850 freezing microtome (Leica, Germany) and placed on Superfrost Plus microscope slides (Thermo, United States). The sections were covered with 1 M HCl and incubated for 1 min, and then, an equal volume of 1% phloroglucinol was added. After 10 min, the frozen sections were observed by a DMi8 inverted fluorescence microscope (Leica, Germany) with a 400× microscopic field.

Data Statistics

The stone cell density of the fruits was calculated as the proportion of the phloroglucinol-HCl stained area to the total area in the equatorial cross-section of the fruit pulp. Statistical analysis was undertaken using ImageJ software. The length and width of stone cells in the microscopic field of the frozen sections were measured. GraphPad Prism 8 (GraphPad Software, United States) was used for data statistics

and plotting. One-way ANOVA was used to test significant differences among three columns, while Student's *t*-test was used for two columns.

RESULTS

Determination of Stone Cell Content in the Common and Wild Loquat Parents and Interspecific Hybrids

As shown in **Table 1**, the stone cell content of the common loquat JFZ was very low (0.03%), while the contents of the wild loquat TL (1.59%) and TLk (0.95%) were at least one order higher than that of JFZ. The interspecific hybrid JT8 (from JFZ×TL) had a high stone cell content, of which the mass ratio was up to 1.61%, higher than that of the male parent TL. In addition, high stone cell contents were also observed in the interspecific hybrids JTk2 and JTk3 (from JFZ×TLk), with mass ratios of 0.79 and 0.96%, respectively. The results showed that the high stone cell contents of the interspecific hybrids probably came from the male parent wild loquat, which could exhibit transgressive inheritance. Microscopic observation of stone cells extracted from the fruits of the interspecific hybrid JT8 revealed many large stone cells or stone cell clusters (**Figure 1**). This may be the main reason for the granular taste of JT8 fruits.

TABLE 1 | Stone cell contents in different loquat plants.

Loquat genotypes	Pulp weight (g)*	Stone cell weight (g)	Proportion (w/w)
JFZ	30.83	0.01	0.03%
TL	32.12	0.51	1.59%
TLk	36.58	0.34	0.95%
JT8	34.79	0.56	1.61%
JTk2	35.25	0.28	0.79%
JTk3	30.11	0.29	0.96%

**The total weight of pulps used to determine stone cell content.*

Stone Cell Distribution Traits in the Common and Wild Loquat Parents and Interspecific Hybrids During Fruit Development

Cross-sections of JFZ, TL, and JT8 fruits with phloroglucinol-HCl staining were observed and are displayed in **Figure 2**. The observation of TL was only performed at four time points at 14, 98, 112, and 126 DAF because of its few fruits. The fruit size of JFZ was significantly larger than those of TL and JT8, and the fruit size of JT8 was between those of the two parents and closer to the male parent TL. At 14 DAF, the outlines of numerous cells were light purple in the fruits of three *Eriobotrya* plants, indicating that lignin deposition and secondary cell wall thickening had occurred and that the formation of stone cells had begun before 14 DAF. After that, the stained cell outline became increasingly thick, the color increasingly deepened, and an increasing number of entire cells became dark purple. The results showed that with the development of fruits, lignin continued to fill in the sclerenchyma cells until the protoplasts disappeared and the stone cells were completely developed. The stone cell density of the JFZ rapidly increased from 5.16% at 14 DAF to 27.58% at 42 DAF, reached a maximum value of 28.40% at 56 DAF, and then slightly decreased to 25.24% at 70 DAF. It then rapidly decreased to 10.68% at 85 DAF, after which it gradually declined to 3.98% until the mature stage at 126 DAF. The stone cells tended to aggregate into small SCCs in the JFZ. The stone cell density of the wild loquat TL was significantly higher than that of JFZ and lower than that of JT8 at the four sampled time points. The stone cells increased from 9.85% at 14 DAF to 26.12% at 98 DAF with a dense and wide distribution and then decreased to 8.25% at 126 DAF, at which time many large stone cell clusters were still apparent. In the interspecific hybrid JT8, the stone cell density also presented a trend of early increase and later decrease. The stone cell density increased from 15.77% at 14 DAF to a maximum value of 31.19% at 98 DAF, followed by a decrease to 19.77% at 126 DAF. The time of peak stone cell density in JT8 may be the same as

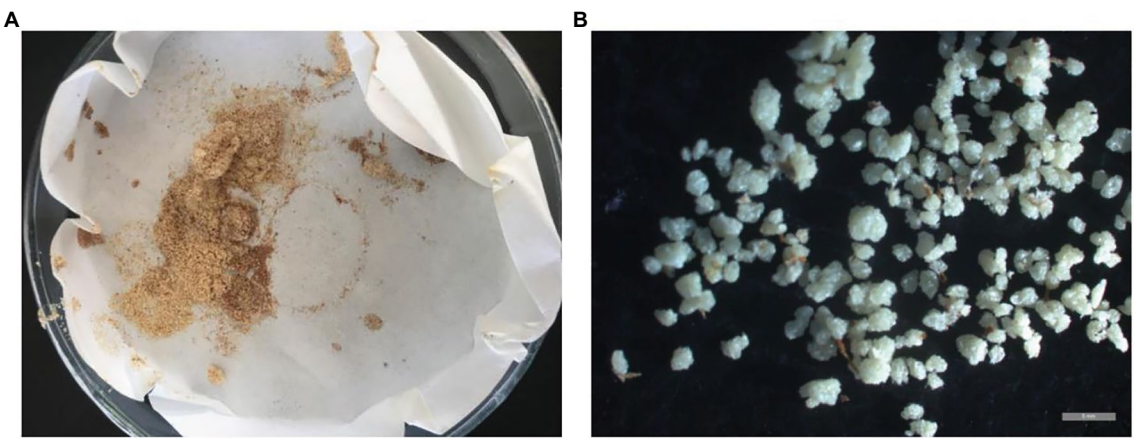


FIGURE 1 | The stone cells extracted from JT8. **(A)** The appearance of stone cells. **(B)** The microscopic morphology of stone cells. The gray bar represents 1 mm.

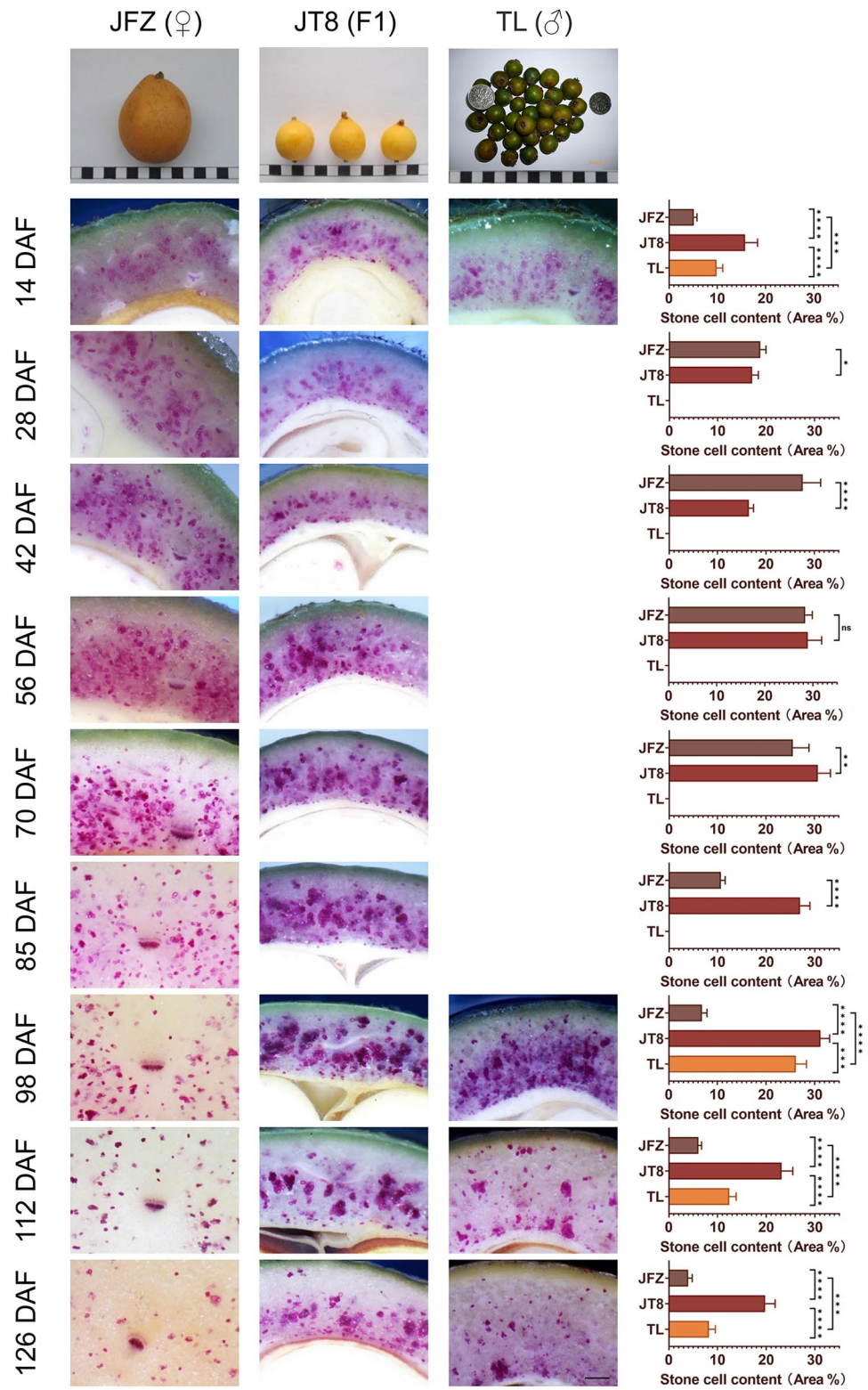


FIGURE 2 | Phloroglucinol-HCl staining and microscopic observation of transverse sections of fruit pulp from three loquat plants. Due to the few samples in wild loquat TL, only four developmental stages were examined, including 14, 98, 112, and 126 DAF. Three fruit images at the top of the figure show the mature fruit size, shape, and color of three loquat plants; in particular, the color of mature TL fruits was brown green. The width of each black or white box represents 1 cm; the black bar at lower right represents 500 μm; bar charts show the stone cell density; and "ns" represents that there is no significant difference, "*" represents value of $p < 0.05$, "**" represents value of $p < 0.01$, "***" represents value of $p < 0.001$, and "****" represents value of $p < 0.0001$.

that of the male parent TL and later than that of the female parent JFZ. The dense distribution of stone cells and many large SCCs also appeared in JT8, similar to the male parent TL. The density and aggregation traits of stone cells of JT8 were similar to those of the male parent TL, suggesting that these stone cell traits were derived from the male parent. Interestingly, the density of stone cells and large SCCs was obviously higher than that of the two parents, indicating that some traits of stone cells may exhibit transgressive inheritance. In fruits at 126 DAF, the density of stone cells was higher near the core than near the peel, and the SCCs near the core were larger than those near the peel, while the distribution of stone cells in TL near the peel was higher than that in JT8 and JFZ, which indicated that some traits of stone cells could be inherited from the female parents. Interestingly, no obvious stone cells could be found in the peel tissues of the three *Eriobotrya* plants. In brief, the key finding was that stone cell traits could be transmitted from parents to progenies in interspecific crosses and that the transgressive inheritance of stone cell traits might occur.

Morphological and Developmental Traits of Stone Cells in Common and Wild Loquat Parents and Interspecific Hybrids

The stone cells of three *Eriobotrya* plants are brachysclereids, and their developmental processes are almost the same. The development of stone cells was accompanied by the thickening of secondary cell walls and the shrinkage of protoplasts until the formation of typical stone cells filled with secondary cell walls occurred, and the protoplasts often disappeared. The pits of stone cells were thicker and clear with branching in the JFZ but thinner and unclear in the TL. There were more pairs of interlinked pits between adjacent stone cells in the JFZ and fewer in the TW. The shape and size of pits in JT8 were similar to those in the female parent JFZ (**Figures 3A–C**). The statistical results showed that the length and width of stone cells and SCCs were not evenly distributed and exhibited great variation. The stone cell length and width of JFZ and JT8 varied widely, while those of TL did not. The median and quartile values of the length and width of stone cells between JFZ and JT8 were similar and were significantly greater than those of TL

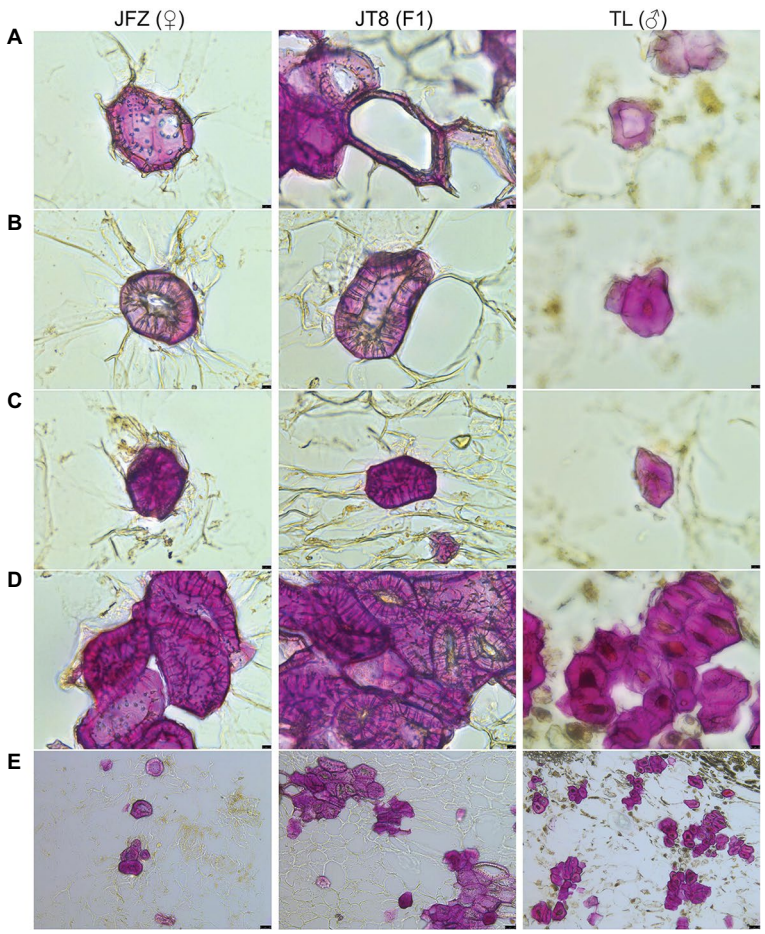


FIGURE 3 | Phloroglucinol-HCl staining and microscopic observation of freezing microtome sections of fruit pulp taken from three loquat plants at 126 DAF. **(A)** The secondary cell wall of sclereid primordium cells starts to thicken (400×). **(B)** The sclereid primordium cell shows a continuously thickened secondary cell wall and shrunken protoplasm (400×). **(C)** The typical stone cell is filled by secondary cell wall structures without protoplasm (400×). **(D)** Stone cell clusters (400×). **(E)** Stone cell clusters (100×). The black bar in the lower right represents 10 μm in **(A,B)** and 50 μm in **(E)**.

(**Figures 4A,B**). The results indicated that the traits of shape and size of stone cells tended to be inherited from the female parent JFZ. In addition, the density of stone cells in the JFZ was sparse, there were only a few stone cells aggregated in single SCCs, while the density of stone cells in the TL was higher, and there were many stone cells aggregated in single SCCs (**Figures 3D,E**). The statistical results showed that the ranges of stone cell length and width of the two parents were small, while those of JT8 were very dispersed. The median and quartile values of the length and width of TL stone cells were greater than those of JFZ stone cells, and both of them were smaller than those of JT8 stone cells. The length of most JT8 SCCs, a small proportion of TL SCCs and some JFZ SCCs, and the width of more than half of JT8 SCCs and a small proportion of TL SCCs were greater than 250 m (**Figures 4C,D**). The results illustrated that although the size of TL stone cells was small, the stronger aggregation of TL stone cells led to larger-sized SCCs in the TL than in the JFZ. However, the aggregation of stone cells in JT8 tended to be derived from the male parent TL, and the size of stone cells tended to be inherited from the female parent JFZ, which resulted in the size of JT8 SCCs exhibiting transgressive inheritance. It is worth noting that in the fruits at 126 DAF, newborn sclereid primordium cells with thin secondary cell walls could be observed in JFZ and JT8 but were rare in TL. This result suggested that JFZ and JT8 fruits may be sustained to produce new stone cells, while TL may be incapable in this regard after a certain time point. These results also supported our previous statement that stone cell traits could be transmitted from parents to progenies and between different species, and that the transgressive inheritance of stone cell traits might occur.

The Distribution of Stone Cells in Two Loquat Cultivars With Different Storage and Transport Tolerances During Fruit Development

BL was susceptible to storage and transport, its pulp texture was very fine, while JFZ was tolerant to storage and transport, and its pulp texture was coarse (Lin et al., 2008). As shown

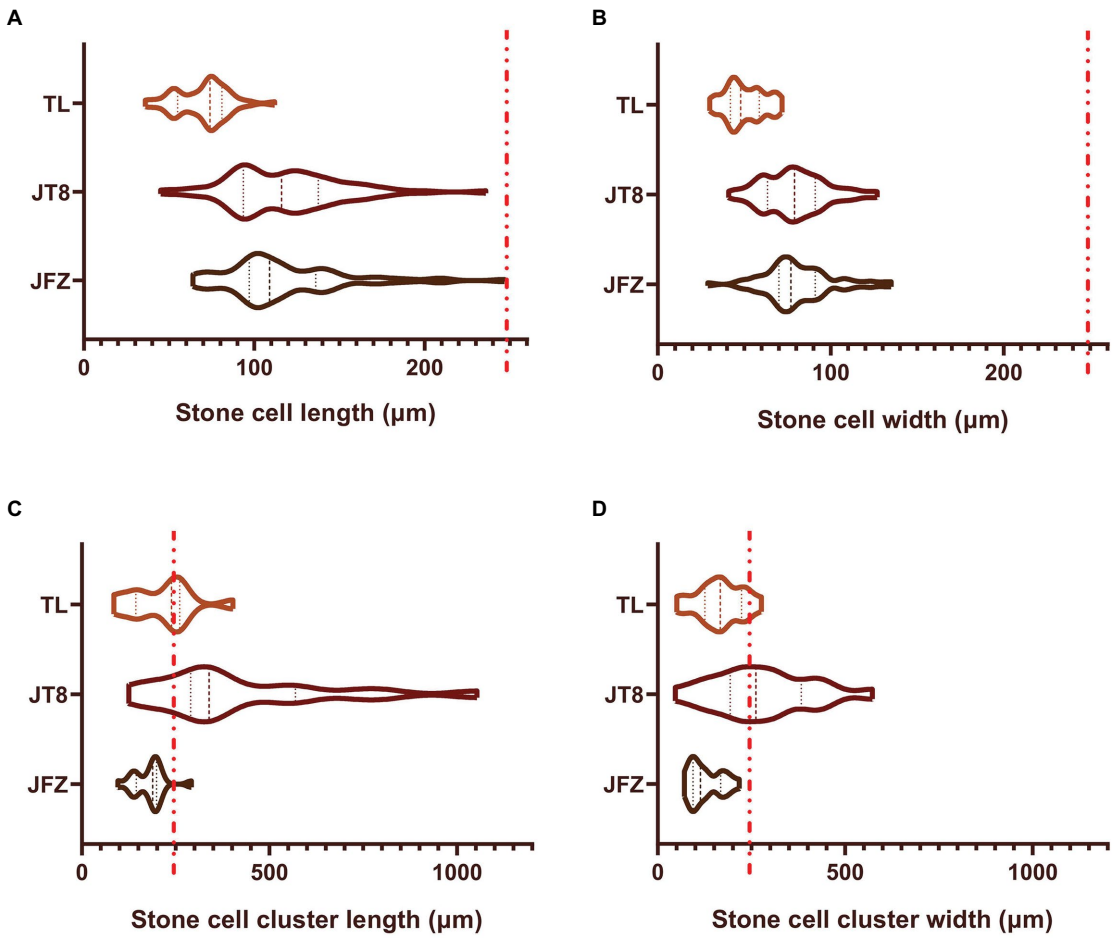


FIGURE 4 | The sizes of stone cells and stone cell clusters in freezing microtome sections of fruit pulp collected at 126 DAF from three loquat plants as determined through phloroglucinol-HCl staining and microscopic observation. Violin plots are used to display the distribution of data sets, and similar violin shapes represent similar data distributions. The thick dashed lines represent the median values, and the two thin dashed lines represent the first and third quartiles. The wider parts of the violin parts correspond to higher probabilities of observed values, while the narrower parts correspond to lower probabilities. Longer and sharper ends of the violin plots correspond with more outliers. The red-dashed line indicates 250 μm. **(A)** Stone cell length; **(B)** Stone cell width; **(C)** Stone cell cluster length; **(D)** Stone cell cluster width.

in **Figure 5**, the stone cell density in BL fruit remained at approximately 1.00% from 14 DAF to 98 DAF and then decreased to 0.32% at 126 DAF. In contrast, the density of stone cells in the JFZ first increased from 5.39% at 14 DAF to 30.67% at 63 DAF and then decreased to 3.51% at 126 DAF. Undoubtedly, the number of stone cells in the JFZ was much higher than that in the BL. The results showed that stone cell traits were correlated with the storage and transport tolerance and pulp texture of loquat fruits and that higher stone cell content and density might be beneficial to enhance the storage and transport tolerance while worsening the taste of loquat fruits. A 20× stereoscopic microscope revealed that many stone cells were scattered in 126 DAF JFZ fruits, and some of them were aggregated into small SCCs, while only a few single stone cells were found in 126 DAF BL fruits. This may be the reason why JFZ has a less desirable taste than BL.

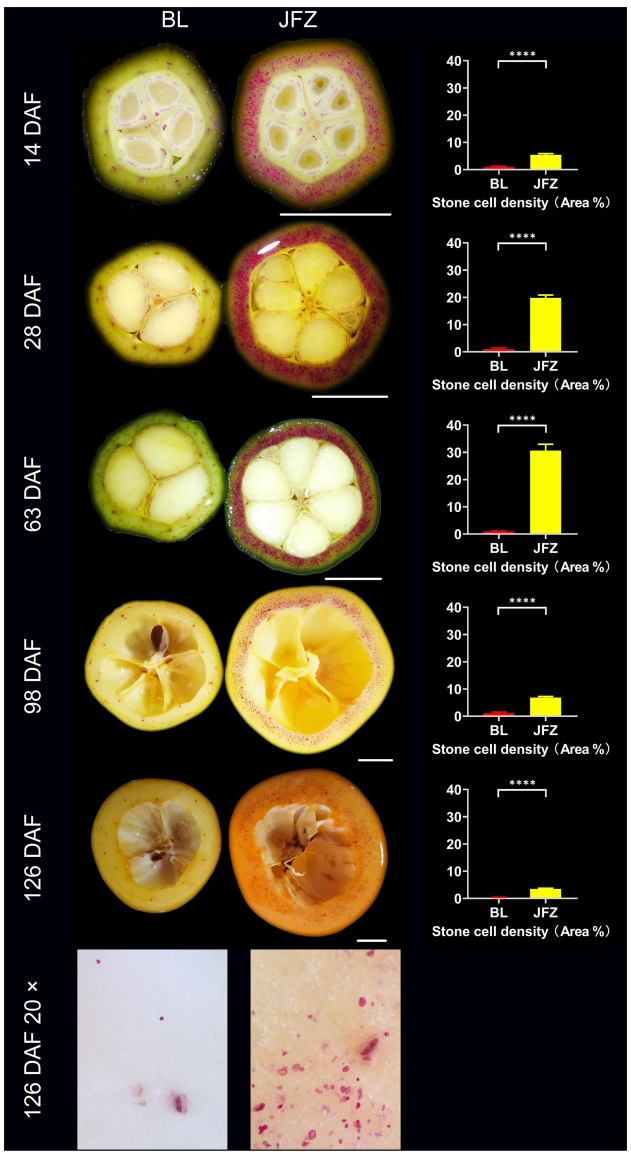


FIGURE 5 | Phloroglucinol-HCl staining and microscopic observation of transverse sections of fruit pulp from white and yellow loquat fruits. DAF: Days after full bloom; BL: White-fleshed loquat fruit (*Eriobotrya japonica* Lindl. cv. Baili); JFZ: Yellow-fleshed loquat fruit (*Eriobotrya japonica* Lindl. cv. Jiefangzhong); and 20×: 20× stereoscopic microscopy observations. Bar charts show the stone cell density; "****" represents value of $p < 0.0001$.

DISCUSSION

Stone cells (or sclereids) are sclerenchyma cells characterized by thickening and lignified cell walls (Nii et al., 2008; Zhang et al., 2017). Stone cells can be classified into five types according to their morphology: brachysclereids, macrosclereids, osteosclereids, astrosclereids, and trichosclereids (Zhao and Zhu, 2014). Stone cells are generally believed to have greater hardness than parenchyma cells and thus could serve a supporting function (Brahem et al., 2017). Moreover, stone cells have also been found to act as a physical defense against white pine weevils in Sitka spruce (Whitehill et al., 2016). At present, most studies on stone cells have been carried out in pear fruits, while only a few have mentioned loquat (Lin et al., 2009; Zhu et al., 2018). In this study, the shape, size, development, and distribution dynamics of fruit stone cells of *Eriobotrya* plants were thoroughly studied. The stone cell traits of *Eriobotrya* plants were highly similar to those of pear, especially the following traits. Pear stone cells are brachysclereids and often include branching pits (Choi and Lee, 2013). Pear stone cells can be found as single cells or as SCCs, and they are usually distributed as SCCs (Wu et al., 2013; Brahem et al., 2017). Some pairs of interlinked pits could often occur between adjacent stone cells. The secondary cell walls of some pear parenchyma cells are thickened first to form sclereid primordium cells, and then the secondary cell walls are continuously thickened until the protoplasm disappears and the cell fills with secondary cell wall structures. The stone cell development is then complete (Jin et al., 2013; Zhao et al., 2013). With the development of pear fruits, the rate of stone cell production was higher than that of fruit expansion, and therefore, the density of stone cells increased until the peak. Subsequently, the rapid expansion of parenchyma cells resulted in a fruit expansion rate greater than the stone cell production rate, resulting in a continuous decrease in stone cell density. The time for stone cells to reach the peak distribution density varied among different pear varieties. In addition, the distribution of stone cells in the pear pulp was not uniform. The density of stone cells was higher and the size of the SCCs larger near the core, but the opposite was found near the peel (Nii et al., 2008; Choi and Lee, 2013; Zhang et al., 2017). These similar traits between loquat and pear stone cells indicated the existence of a relationship between loquat and pear, which may provide new ideas for the study of the relationship and evolution among the fruit trees of *Rosaceae*. Moreover, due to the highly similar features of stone cell traits between loquat and pear fruits, the relevant research results of pear stone cells can be used as an important reference for the study of loquat stone cells.

The pear-like granular taste is hardly found in the fruit of loquat cultivars, so there is not much consumer concern about whether loquat fruits contain stone cells. Prof. Lin Shunquan, one of the authors of this paper, collected 26 species of *Eriobotrya* and used seven species of *Eriobotrya* and two- related species to conduct 91 cross combinations between *Eriobotrya* plants and 21 cross combinations between *Eriobotrya* plants and related plants from 2004 to 2014 (Lin, 2017). Previous studies found that the fruit of the interspecific hybrid JT8 obviously had a granular taste similar to that of pear fruits (unpublished). In this study, we found that some stone cell traits were significantly different among *Eriobotrya* plants. The key finding of this study is that stone cell traits could be transmitted from parents to progenies in interspecific crosses, and the transgressive inheritance of stone cell traits might occur. Therefore, it is feasible to transfer stone cell traits of wild loquat into common loquat cultivars by cross-breeding. The content of pear stone cells exhibited quantitative trait inheritance, and the content of pear stone cells in hybrid offspring showed a trend of significant increase compared with their parents (Cui et al., 2011; Zhang et al., 2018). It has also been reported that the content of stone cells in hybrid offspring tended to be distributed around the mid-parent value, and the contents of stone cells were greatly affected by the male parent (Bai et al., 2018). The size and density of pear SCCs also showed quantitative trait inheritance, and the size and density of the hybrid offspring tended to be larger and smaller than those of the parent, respectively. The degree of variation was different from those of the parental varieties (Cui et al., 2011; Zhang et al., 2018). The propensity for the inheritance of pear stone cells could be used to partly explain the stone cell traits of the interspecific hybrid JT8 in this study, but determining whether the inheritance of stone cell traits in the loquat interspecific hybrid population is similar to that of pear requires further study.

Compared with apple, pear, and many other *Eriobotrya* fruits, loquat fruits have poor storage and transport tolerance, which is one of the problems that the industry urgently needs to solve (Lin et al., 2018). A few reports have revealed a certain correlation between stone cell traits and storage and transport tolerance in loquat cultivars (Lin et al., 2009; Zhu et al., 2018). We further verified the relationship between stone cell traits and storage and transport tolerance in loquat fruits and found that there were few stone cells during the whole fruit development process in a storage- and transport-susceptible cultivar. This suggested that stone cell traits have potential application value to improve the storage and transport tolerance of loquat cultivars. However, all of the main loquat cultivars were domesticated from a species of *Eriobotrya* named common loquat, whose genetic basis was very narrow (Wang and Lin, 2012). In contrast, the genetic basis of pear cultivars is much broader, and at least five species of pear plants are ancestors of domesticated cultivars, including *Pyrus ussuriensis*, *P. sinkiangensis*, *P. pyrifolia*, *P. communis*, and *P.* x *bretschneideri*. The stone cell traits of these cultivars and their interspecific hybrids showed abundant diversity (Cao et al., 2010; Tian et al., 2011). Stone cell traits, such as content, size, and density, can affect the quality of pear fruits, so one of the main goals of pear breeding is to breed excellent varieties with fewer stone cells (Choi and Lee, 2013; Zhang et al., 2017, 2018). Nevertheless, the results of this study showed that the content and density of stone cells and the diameter of SCCs in loquat cultivars were relatively small. Therefore, it was necessary to use interspecific hybridization to transfer appropriate stone cell traits from wild loquat into loquat cultivars, thus improving the storage and transport tolerance of loquat cultivars.

JFZ was tolerant to storage and transport, while BL was susceptible to these factors. However, the pulp texture was coarse in the JFZ and very fine in the BL (Lin et al., 2008). Both JFZ and BL are regarded as "standard" varieties of "fruit hardness" and other fruit traits in the Chinese DUS test standard of loquat (Lin et al., 2019). In China, breeding research attention has traditionally been placed on maintaining the original taste of loquat by ensuring a high TSS content, with storage and transport tolerance only being considered in recent years. Our results suggested that higher stone cell content and density might be beneficial to enhance storage and transport tolerance while worsening the taste of loquat fruits. Therefore, it will be necessary to investigate the correlations between storage and transport tolerance and fruit texture and taste in loquat to ensure a good balance between these characteristics in future breeding programs.

Different pear varieties exhibit diverse stone cell traits. A large diameter and density of SCCs lead to a more crude and granular taste. The effect of SCCs with diameter greater than 250 μm on pulp texture is significant (Cao et al., 2010; Tian et al., 2011). In this study, although the size of TL stone cells was small, the stronger aggregation of TL stone cells led to a larger size of SCCs in this species than in JFZ. However, JT8 SCCs combined the aggregation traits from the male parent TL and the size traits from the female parent JFZ, thereby exhibiting transgressive inheritance for the SCC size-related traits. The length and width of most JFZ SCCs were less than 250 μm. In contrast, the length of most SCCs and the width of more than half of the SCCs were greater than 250 μm in JT8, and certain SCCs could exceed 1,000 μm in length and width. This may be the reason why JT8 fruits had an obviously granular taste, while JFZ fruits did not. Therefore, efforts to improve the storage and transport tolerance of loquat cultivars by using stone cell traits from wild loquat should fully consider the effects of stone cell traits on the taste of interspecific hybrid fruits. At the same time, a common disadvantage of interspecific hybrids is that their fruits are usually smaller than those of loquat cultivars (e.g., JFZ); therefore, backcrossing is essential (Lin, 2017). In addition, it is worth noting that there was high stone cell content in the peels of pear fruits, and the distribution density of stone cells was as follows: within the peel > in the pulp near the core > in the pulp near the peel (Zhang et al., 2017). Stone cells have high mechanical strength and could fulfill a supporting function and even a role in physical defense against insect herbivory (Whitehill et al., 2016). In this study, no obvious stone cells could be found in the peel tissues of three *Eriobotrya* plants, which we considered to be one of the important reasons why loquat fruits were susceptible to storage and

transport tolerance. It is not clear whether there are wild loquats with stone cells distributed in the peels. Therefore, the fruit stone cell traits in wild loquat need to be further studied. Wild loquat with desirable stone cell traits (e.g., content and size or distribution in fruit peels) will be beneficial for improving the storage and transport tolerance of loquat cultivars by using these traits.

CONCLUSION

Overall, the shape, size, development, and distribution dynamics of stone cells of *Eriobotrya* plants were thoroughly studied. It was found that the stone cell traits of *Eriobotrya* plants were highly similar to those of pear, indicating a relationship between loquat and pear, which may provide a new idea for the study of the relationship and evolution among the fruit trees of *Rosaceae*. Moreover, the results also demonstrated that stone cell traits could be transmitted from parents to progenies in interspecific crosses. Thus, it is feasible to transfer stone cell traits of wild loquat into common loquat cultivars by cross-breeding. Our results provide a new approach to improving the storage and transport tolerance of loquat cultivars through the use of stone cell traits from wild loquat plants.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JW designed and supervised the experiment. SL and DL mainly performed the research and drafted the manuscript. BW and SM carried out the statistical data analysis. SS, TZ, WZ, YB, and QW finished specific parts of the experiments. JW and SL revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the Natural Science Foundation of Fujian Province (2019J01809 and 2021J05240), the Fujian Provincial Science and Technology Project (2021N5014), the Education and Research Project of Young and Middle-Aged Teachers of Fujian Province (JAT170501 and JAT200524), the Research Project of Putian Science and Technology Bureau (2018ZP08 and 2020NP001), the Research and Innovation Special Foundation of Putian University (2016CX001), and the Scientific Research Project of Putian University (2016069 and 2018064).

ACKNOWLEDGMENTS

We thank Prof. Shunquan Lin for his help and advice, especially providing us with fruit samples of wild loquat and interspecific hybrids. This manuscript was edited by American Journal Experts (AJE) for proper English language use.

REFERENCES

Bai, B., Liu, L., Qin, M., Gu, C., and Wu, J. (2018). Studies on genetic tendency of fruit characters in reciprocal crosses generation between 'Dangshansuli' and'Fengshuili'pear cultivars. *J. Jiang. Agric. Sci.* 46, 108–112. doi: 10.15889/j.issn.1002-1302.2018.16.027

Brahem, M., Renard, C. M., Gouble, B., Bureau, S., and Le Bourvellec, C. (2017). Characterization of tissue specific differences in cell wall polysaccharides of ripe and overripe pear fruit. *Carbohydr. Polym.* 156, 152–164. doi: 10.1016/j.carbpol.2016.09.019

Cao, Y.-F., Tian, L.-M., Li, L.-L., and Gao, Y. (2010). Comparison studies on the stone cell content in flesh of pear cultivars. *Acta Hortic. Sin.* 37, 1220–1226. doi: 10.16420/j.issn.0513-353x.2010.08.016

Choi, J.-H., Choi, J.-J., Hong, K.-H., Kim, W.-S., and Lee, S.-H. (2007). Cultivar differences of stone cells in pear flesh and their effects on fruit quality. *Hortic. Environ. Biotechnol.* 48, 27–31.

Choi, J.-H., and Lee, S.-H. (2013). Distribution of stone cell in Asian, Chinese, and European pear fruit and its morphological changes. *J. Appl. Bot. Food Qual.* 86, 185–189. doi: 10.5073/JABFQ.2013.086.025

Cui, Y., Chen, H., Yue, W., Zhang, S., Wu, T., Tao, S., et al. (2011). Studies on genetic tendency of fruit characters in reciprocal crosses generation between'Jingbaili'and'Yali'pear cultivars. *Acta Hortic. Sin.* 38, 215–224. doi: 10.16420/j.issn.0513-353x.2011.02.004

Huang, Z., Mao, Q.-Q., and Liu, M.-L. (2005). 18 kinds of Chinese traditional Medicine's sclerotic cells microcinematography. *J. Wenzhou Teach. College* 26, 80–85.

Huang, W., Zhu, N., Zhu, C., Wu, D., and Chen, K. (2019). Morphology and cell wall composition changes in lignified cells from loquat fruit during postharvest storage. *Postharvest Biol. Technol.* 157:110975. doi: 10.1016/j.postharvbio.2019.110975

Jin, Q., Yan, C., Qiu, J., Zhang, N., Lin, Y., and Cai, Y. (2013). Structural characterization and deposition of stone cell lignin in Dangshan Su pear. *Sci. Hortic.* 155, 123–130. doi: 10.1016/j.scienta.2013.03.020

Lin, S. (2017). A review on research of the wild species in genus eriobotrya germplasm and their innovative utilization. *Acta Hortic. Sin.* 44, 1704–1706. doi: 10.16420/j.issn.0513-353x.2017-0488

Lin, J., Lin, H., Guo, Z., and Chen, S. (2009). Morphology and structure and their relationships to storability of loquat fruits (*Eriobotrya japonica* Lindl.). *Chin. J. Trop. Crops* 30, 53–58.

Lin, J., Lin, H., Huang, Z., Chen, G., and Zheng, Y. (2008). Comparison of fruit qualities of five major loquat cultivars in Fujian Province and relationship between fruit quality and storability. *Food Sci.* 29, 433–437.

Lin, S., Rao, D., Dai, Y., Huang, B., Gao, Y., Yang, X., et al. (2019). Guidelines for the conduct of tests for distinctness, uniformity and stability--Loquat (*Eriobotrya* Lindl.). NY/T 3433–2019.

Lin, S., Wu, T., Lin, H., Zhang, Y., Xu, S., Wang, J., et al. (2018). De novo analysis reveals transcriptomic responses in *Eriobotrya japonica* fruits during postharvest cold storage. *Genes* 9:639. doi: 10.3390/genes9120639

Liu, X., Li, S., Feng, X., and Li, L. (2020). Study on cell wall composition, fruit quality and tissue structure of hardened 'Suli'Pears (*Pyrus bretschneideri* Rehd). *J. Plant Growth Regul.* 40, 1–10. doi: 10.1007/s00344-020-10248-4

Nie, J., Li, J., Yang, Z., Zhang, H., and Li, M. (2006). Study on the conditions for measuring stone cell content in pear flesh by freezing method. *J. Fruit Sci.* 23, 133–135.

Nii, N., Kawahara, T., and Nakao, Y. (2008). The development of stone cells in Japanese pear fruit. *J. Hortic. Sci. Biotechnol.* 83, 148–153. doi: 10.1080/14620316.2008.11512362

Pan, Z. (2011). Zhongtian Dashanzha, a new hawthorn cultivar for fresh market in the south of China. *J. Fruit Sci.* 28, 186–187. doi: 10.13925/j.cnki.gsxb.2011.01.031

Sun, Y.-Y., and Li, B.-S. (2006). The traditional Chinese medicine varieties were identified by stone cells. *Shizhen Guo Yi Guo Yao* 17, 1256–1257.

Tao, S., Khanizadeh, S., Zhang, H., and Zhang, S. (2009). Anatomy, ultrastructure and lignin distribution of stone cells in two Pyrus species. *Plant Sci.* 176, 413–419. doi: 10.1016/j.plantsci.2008.12.011

Tian, L., Cao, Y., Gao, Y., and Dong, X. (2011). Effect of stone cells size and flesh texture in pear cultivars. *Acta Hortic. Sin.* 38, 1225–1234. doi: 10.16420/j.issn.0513-353x.2011.07.003

Wang, Y., and Lin, S. (2012). A comparative study on the polymorphism of loci cpDNA~(TrnS-TrnG) and cpDNA~(TrnQ-rps16) in cultivated and wild loquats. *Acta Hortic. Sin.* 39, 1913–1918. doi: 10.16420/j.issn.0513-353x.2012.10.012

Whitehill, J. G. A., Henderson, H., Strong, W., Jaquish, B., and Bohlmann, J. (2016). Function of Sitka spruce stone cells as a physical defence against white pine weevil. *Plant Cell Environ.* 39, 2545–2556. doi: 10.1111/pce.12810

Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhu, S., et al. (2013). The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res.* 23, 396–408. doi: 10.1101/gr.144311.112

Zhang, J., Cheng, X., Jin, Q., Su, X., Li, M., Yan, C., et al. (2017). Comparison of the transcriptomic analysis between two Chinese white pear (*Pyrus bretschneideri* Rehd.) genotypes of different stone cells contents. *PLoS One* 12:e0187114. doi: 10.1371/journal.pone.0190180

Zhang, Q., Jiang, X., Duan, H., and Li, J. (2018). Studies on the characteristics and genetic tendency of the stone cell in hybrid offspring of'Kuerlexiangli'pear. *J. Fruit Sci.* 35, 89–96. doi: 10.13925/j.cnki.gsxb.2018.S.14

Zhao, S.-G., Zhang, J.-G., Zhao, Y.-P., and Zhang, Y.-X. (2013). New discoveries of stone cell differentiation in fruitlets of 'Yali'pears (Pyrus bretschneideri Rehd.). *J. Food Agric. Environ.* 11, 937–942.

Zhao, M., and Zhu, H. (2014). Development and morphology of stone cells in phloem of Toxicodendron vernicifluum. *Trees* 28, 1553–1558. doi: 10.1007/s00468-014-1027-9

Zhu, N., Wu, D., and Chen, K. (2018). Label-free visualization of fruit lignification: Raman molecular imaging of loquat lignified cells. *Plant Methods* 14:58. doi: 10.1186/s13007-018-0328-1

Zhu, N., Yang, Y., Ji, M., Wu, D., and Chen, K. (2019). Label-free visualization of lignin deposition in loquats using complementary stimulated and spontaneous Raman microscopy. *Hortic. Res.* 6:71. doi: 10.1038/s41438-019-0153-3

Zhu, N., Zhao, C., Wei, Y., Sun, C., Wu, D., and Chen, K. (2021). Biosynthetic labeling with 3-O-propargylcaffeyl alcohol reveals in vivo cell-specific patterned lignification in loquat fruits during development and postharvest storage. *Hortic. Res.* 8:61. doi: 10.1038/s41438-021-00497-z

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fpls.2021.813798

Comparative Analysis of the MADS-Box Genes Revealed Their Potential Functions for Flower and Fruit Development in Longan (*Dimocarpus longan*)

Baiyu Wang[1†], Wenshun Hu[1,2†], Yaxue Fang[1], Xiaoxi Feng[1], Jingping Fang[3], Tengyue Zou[4], Shaoquan Zheng[2], Ray Ming[5] and Jisen Zhang[1]*

[1] Center for Genomics and Biotechnology, Haixia Institute of Science and Technology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, College of Life Sciences, Fujian Agriculture and Forestry University, Fuzhou, China, [2] Fujian Breeding Engineering Technology Research Center for Longan & Loquat, Fruit Research Institute, Fujian Academy of Agricultural Sciences, Fuzhou, China, [3] College of Life Sciences, Fujian Normal University, Fuzhou, China, [4] College of Mechanical and Electrical Engineering, Fujian Agriculture and Forestry University, Fuzhou, China, [5] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, United States

Longan (*Dimocarpus longan* Lour.) is an important economic crop widely planted in tropical and subtropical regions, and flower and fruit development play decisive effects on the longan yield and fruit quality formation. MCM1, AGAMOUS, DEFICIENS, Serum Response Factor (MADS)-box transcription factor family plays important roles for the flowering time, floral organ identity, and fruit development in plants. However, there is no systematic information of MADS-box family in longan. In this study, 114 MADS-box genes were identified from the longan genome, phylogenetic analysis divided them into type I ($M\alpha$, $M\beta$, $M\gamma$) and type II (*MIKC**, $MIKC^C$) groups, and $MIKC^C$ genes were further clustered into 12 subfamilies. Comparative genomic analysis of 12 representative plant species revealed the conservation of type II in Sapindaceae and analysis of cis-elements revealed that *Dof* transcription factors might directly regulate the $MIKC^C$ genes. An ABCDE model was proposed for longan based on the phylogenetic analysis and expression patterns of MADS-box genes. Transcriptome analysis revealed that $MIKC^C$ genes showed wide expression spectrums, particularly in reproductive organs. From 35 days after $KClO_3$ treatment, 11 *MIKC* genes were up-regulated, suggesting a crucial role in off-season flower induction, while *DlFLC*, *DlSOC1*, *DlSVP*, and *DlSVP-LIKE* may act as the inhibitors. The gene expression patterns of longan fruit development indicated that *DlSTK*, *DlSEP1/2*, and *DlMADS53* could be involved in fruit growth and ripening. This paper carried out the whole genome identification and analysis of the longan MADS-box family for the first time, which provides new insights for further understanding its function in flowers and fruit.

Keywords: longan, MADS-box, ABCDE model, $KClO_3$, flower, fruit development

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Hui Song,
Qingdao Agricultural University, China
Qilin Tian,
Minnan Normal University, China
Lisha Shen,
Temasek Life Sciences Laboratory, Singapore

****Correspondence:***
Jisen Zhang
zjisen@fafu.edu.cn

† These authors have contributed equally to this work and share first authorship

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

Received: *12 November 2021*
Accepted: *16 December 2021*
Published: *27 January 2022*

Citation:
Wang B, Hu W, Fang Y, Feng X, Fang J, Zou T, Zheng S, Ming R and Zhang J (2022) Comparative Analysis of the MADS-Box Genes Revealed Their Potential Functions for Flower and Fruit Development in Longan (Dimocarpus longan). Front. Plant Sci. 12:813798. doi: 10.3389/fpls.2021.813798

INTRODUCTION

The MADS-box genes, widely distributed in fungi, plants, and animals, encode a large transcription factor family. These genes with diverse functions play important roles in plant development, signal transduction, and stress responses (Riechmann and Meyerowitz, 1997; Messenguy and Dubois, 2003). The family name "MADS" consists of the initials of its earliest genes, *MCM1* in yeast, *AGAMOUS* (*AG*) in *Arabidopsis thaliana* (Arabidopsis), *DEFICIEN*S in *Antirrhinum majus,* and *Serum Response Factor* (*SRF*) in humans (Passmore et al., 1988; Sommer et al., 1990; Yanofsky et al., 1990). Each MADS protein possesses a conserved domain of about 60 amino acids named MADS-box at the N-terminus, which recognizes and binds to CArG boxes (CC[A/T]6GG) (De Bodt et al., 2003). At an evolutionary level, all the members can be divided into two lineages, type I and type II, according to their relationships to animal *SRF-like* and *MEF2-like* genes, respectively (Alvarez-Buylla et al., 2000). In general, type I MADS genes have simple structures with 1–2 exons and there are few reports on their functions. The type II genes were also named *MIKC* genes due to their four domain structures: the highly conserved MADS-box (M) domain, the less-conserved intervening (I) domain, the moderately conserved keratin-like (K) domain, and the variable C-terminal (C) region (Theissen et al., 1996; Kaufmann et al., 2005). In Arabidopsis MADS-box genes were separated into five groups including *M*α, *M*β, *M*γ, *M*δ, and *MIKC* based on the phylogenetic analysis (Parenicova et al., 2003). These *MIKC* type transcription factors, only found in the plant, can be subdivided into $MIKC^C$ and *MIKC** groups based on differences in I and K domains and have been characterized functionally in the regulation of growth and development (Theissen et al., 2000; Becker and Theissen, 2003). These type II genes may be separated from the ancestor of extant land plants (Becker and Theissen, 2003).

In seed plants, extensive research has been performed since the $MIKC^C$ type genes were demonstrated to be involved in floral organ identity, the control of flowering time, and seed development. They can be further classified into 12 clades (Becker and Theissen, 2003), for example, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) and *SHORT VEGETATIVE PHASE* (*SVP*) participate in the regulation of flowering time; *APETALA 1* (*AP1*) is not only a floral meristem identity gene, but also the floral organ identity gene; *FRUITFULL* (*FUL*) regulates cell differentiation during fruit development (Coen and Meyerowitz, 1991; Gu et al., 1998; Lee and Lee, 2010). With extensive researches, the well-known ABCDE model was built from studies in Arabidopsis and *Antirrhinum majus*, which explains floral organ identity (Coen and Meyerowitz, 1991; Theissen, 2001; Theissen and Saedler, 2001). Within the floral meristem, A + E genes specify sepals, A + B + E genes specify petals, B + C + E genes determine stamens, and C + E genes direct carpels, in addition, D + E genes are involved in ovule development. In the model plant Arabidopsis, A-class genes are represented by *AP1*, *APETALA2* (*AP2*), *FUL*, *CAULIFLOWER* (*CAL*), and *AGAMOUS-LIKE 79* (*AGL79*) (Jofuku et al., 1994), B class correspond to genes from *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) (Goto and Meyerowitz, 1994; Yang et al., 2003), C class includes *AG* (Yanofsky et al., 1990), D class includes *SHATTERPROOF* (*SHP*) and *SEEDSTICK* (*STK*) and E class include 4 *SEPALLATA* genes (*SEP1*, *SEP2*, *SEP3*, and *SEP4*) (Pelaz et al., 2000; Pinyopich et al., 2003; Ditta et al., 2004; Zahn et al., 2006).

Longan (*Dimocarpus longan* Lour., 2n = 2x = 30), a member of the Sapindaceae family, was derived and widely cultivated in Southeast Asia (Lai et al., 2000). As a famous subtropical fruit tree, longan commonly called "dragon eye" in China is of high nutritional and medicinal value (Mei et al., 2014). Floral induction (FI) is the biological basis for the development of floral organs and fruit. However, in longan, FI is sensitive to temperature and thus floral reversion occurs frequently (Chen et al., 2009). Began in the late 1990s, potassium chlorate ($KClO_3$) has been heavily input into longan thanks to the discovery of flowering induction by $KClO_3$ at almost any time of the year (Subhadrabandhu and Yapwattanaphun, 2001; Manochai et al., 2005). Interestingly, off-season FI by $KClO_3$ was only found in longan, even its close species litchi (*Litchi chinensis* Sonn.) has never been affected by $KClO_3$. Several studies have investigated the physiological mechanisms of FI by $KClO_3$, indicating that starch, sucrose contents, and cytokinin may play an important role (Sringarm et al., 2009; Chang, 2010). In addition, "Sijimi" ("SJ") longan, performs a unique trait of perpetual flowering (PF), flower and bears fruit throughout the year without external environment conditions (Zhang H.N. et al., 2016). Longan likely has a special FI mechanism among the perennial fruit species. In this decade, the molecular genetics of flowering in model plants has a great advanced development (Fornara et al., 2010), but the research progress of flowering regulation in perennial fruit trees is very limited due to their long juvenile period and complex genetic backgrounds. Several flowering genes have been cloned and analyzed in longans such as *FLOWERING LOCUS T* (*FT*), *AP1,* and *LEAFY* (*LFY*) (Winterhagen et al., 2013), however, no research has been carried out on the MADS-box family in longan.

Recently, we sequenced and assembled the "Shixia" ("SX") longan genome, this chromosome level genome can assist us with the analysis of MADS-box genes from the entire genome. Here, we identified 114 MADS-box genes from the longan genome for the first time, with the phylogenetics, gene evolution, conserved motif, and cis-element analysis was performed. We also determined the expression profiles of longan MADS-box genes particularly the ABCDE model genes during off-season FI under $KClO_3$ and fruit development. This work provides an overview and information useful for future functional analysis of longan MADS-box genes in the reproductive process.

MATERIALS AND METHODS

Data Retrieval and Plant Materials

The Arabidopsis MADS-box genes were retrieved from The Arabidopsis Information Resource (TAIR)[1]. The MADS-box genes of *Amborella trichopoda*, *Citrus clementina*, *Citrus sinensis*, *Vitis vinifera*, *Oryza sativa*, *Sorghum bicolor*, *Zea mays*, *Musa*

[1]http://www.arabidopsis.org/

acuminate, *Nymphaea colorata,* and *Physcomitrella patens* were described in previous reports (Aono and Hasebe, 2006; Arora et al., 2007; Zhao et al., 2011; Hou et al., 2013; Project, 2013; Wang et al., 2015; Liu et al., 2017; Zhang et al., 2020). The genomic data of *Selaginella moellendorffii* and *Chlamydomonas reinhardtii* were obtained from Phytozome[2] and *Xanthoceras sorbifolium* was downloaded from GigaDB[3]. The evolutionary relationships of the above species were obtained from National Center for Biotechnology Information (NCBI)[4]. The RNA-seq data from ten tissues (root, stem, leaf, dormant bud, flower bud, flower, young fruit, pulp, pericarp, and seed) of "Sijimi" ("SJ") longan were downloaded from the NCBI under accession number: PRJNA329283 and PRJNA387674 from previous studies (Lin et al., 2017; Jue et al., 2019).

The collection of *Dimocarpus longan* and the performance of experimental research on such plants complied with the national guidelines of China. We collected apical buds under CK and $KClO_3$ (99% active ingredient) treatment from 4∼6 years old "Shixia" longan in Maoming, Guangdong, China at ten-time points (from day 0 to day 54 including November 18, 2016, November 23, 2016, November 28, 2016; December 3, 2016, December 8, 2016, December 13, 2016, December18, 2016, December 23, 2016, December 29, 2016, and January 11, 2017). Four whorls of floral organs including sepal, petal, stamen, carpel (pistil without ovary) and ovary of "Baoshi No. 1" ("BS-1") and fruit of n were collected in Germplasm Repository of Longan (*Dimocarpus longan*), Fuzhou City, Ministry of Agriculture. All samples were collected and stored at −80°C until RNA-Seq. A pair-end library was made using the Illumina® TruSeq™ RNA Sample Preparation Kit [RS-122-2001(2), Illumina] and sequencing using Illumina X ten.

Identification and Features of MADS-Box Family in Longan

Two strategies were used to identify the MADS-box transcription factor family: the hidden Markov model (HMM) profile of SRF-TF domains (PF00319) from the Pfam database[5] was used as a query to identify MADS-box sequences with HMMER version 3 (Eddy, 2011) against the longan genome with a threshold of e-value $\leq 1 \times 10^{-5}$. In addition, MADS protein sequences of Arabidopsis were used as queries to search against the longan genome using the BLASTP program (Altschul et al., 1990) with an e-value cutoff of 1e−5 and identity > 40%. Subsequently, these proteins were submitted to the Pfam database and NCBI Conserved Domain Search[6] to confirm the presence and completeness of the MADS domain. Candidate genes without the MADS domain were re-annotated manually with the assistance of FGENESH (Solovyev et al., 2006). The physical-chemical properties of longan MADS-box genes were predicted with the online tool ProtParam from ExPASy (see text footnote 6).

Phylogenetic Analysis

Sequence alignments of MADS-box genes were performed using the MUSCLE program in MEGA X (Kumar et al., 2018) with default parameters. The NJ (neighbor-joining) phylogenetic tree was constructed using MEGA X with the following parameters: Poisson model and pairwise deletion, bootstrap for 1000 replicates. Maximum-likelihood (ML) phylogenetic trees were constructed using FastTree software (Price et al., 2009) with the LG model. Arabidopsis MADS-box genes *(AtMADS)* were used to assist classification (Theissen et al., 2000).

Chromosome Locations and Synteny Analysis

To analyze the synteny for longan MADS-box genes (*DlMADS*), the BLASTP program (e-value < 1e−5), and MCScanX (Wang et al., 2012) were used. The tool 'duplicate gene classifier' was used to classify the origins of duplicate genes for longan.

Gene Structure, Conserved Motif, and Cis-Element Analysis

Gene structures of MADS-box genes were extracted from the General Feature Format (GFF) file and the diagram was drawn with the online program Gene Structure Display Server (GSDS[7]). Conserved motifs were identified using Multiple EM for Motif Elicitation (MEME, version 5.3[8]) with the following parameters: 10 different motifs, Minimum Motif Width 10, Maximum Motif Width 100. The identified motifs were sent to Interpro[9] for annotation.

The 1,500-bp sequences upstream of the start codon of each longan MADS-box gene were extracted to predict cis-acting elements through PlantCARE[10]. We further detected the conserved motifs from promoters of *MIKC*^{C} genes using MEME with *p*-values < 0.05 and compared the motifs with known transcription factor binding sites (TFBS) from JASPAR Core plants 2018 database[11] by performing Motif Comparison Tool TOMTOM (TOMTOM version 5.3[12]) with e-value < 0.05.

RNA-Seq Data Analysis

Raw reads from RNA-seq were trimmed with Trimmomatic version 0.39 (Bolger et al., 2014) to remove adaptor sequences and low-quality reads. The high-quality reads were mapped to the "SX" longan genome using HISAT2 version 2.1 (Kim et al., 2015). The expression levels of MADS-box genes were normalized to fragments per kilobase of exons per million fragments mapped (FPKM) using StringTie version 2.1.2 (Kovaka et al., 2019). We defined genes with an FPKM value < 1 as not expressed, genes with FPKM value > 1 as lowly expressed, and values > 10 as highly expressed. Genes with FPKM value > 100 were extremely highly expressed. The differentially expressed genes (DEGs) were

[2]https://phytozome.jgi.doe.gov/pz/portal.html

[3]http://gigadb.org/

[4]https://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi

[5]http://Pfam.sanger.ac.uk/

[6]https://web.expasy.org/protparam/

[7]http://gsds.gao-lab.org/

[8]http://meme-suite.org/tools/meme

[9]https://www.ebi.ac.uk/interpro/search/sequence/

[10]http://bioinformatics.psb.ugent.be/webtools/plantcare/html/

[11]http://jaspar.genereg.net

[12]https://meme-suite.org/meme/tools/tomtom

identified using the R package DESeq2 version 1.3 (Love et al., 2014). The gene co-expression networks were constructed using the WGCNA (Langfelder and Horvath, 2008) package with the filtered genes (FPKM > 1 at least one sample). Software Cytoscape version 3.7.1 (Shannon et al., 2003) was used to visualize the gene interaction networks of MADS-box genes.

Expression Validation of *DlMADS* by qRT-PCR

The cDNA for qRT-PCR was synthesized using the StarScript II First-strand cDNA Synthesis Mix with gDNA Remover (GenStar, A224-10). Gene-specific primers were designed using the online tool PrimerQuest[13] (**Supplementary Table 1**). Two longan actin genes (*Dil.10g021740.1* and *Dil.06g016430.1*) were used as the internal control and three replicates were performed (Jue et al., 2019). The qRT-PCR amplification was carried out using 2× RealStar Green Fast Mixture (GenStar, A301-10) on a Multicolor Real-Time PCR Detection System (Bio-Rad) using the protocol for this kit: 95°C for 2 min, 40 cycles of 95°C for 15 s and 60°C for 30 s. The relative expression levels of the candidate genes were calculated using the $2-\Delta\Delta Ct$ method.

RESULTS

Identification of MADS-Box Genes in Longan

To identify the MADS-box gene family, both the hidden Markov model (HMM) profile (PF00319) and 107 Arabidopsis MADS-box protein sequences were used as queries to perform HMMER and BLASTP against the "SX" longan genome. A total of 117 candidate MADS-box genes were identified. Among them, nine genes that did not have MADS domain-coding sequences were manually re-annotated with the assistance of the online tool FGENESH. Three re-annotated sequences without MADS domain were excluded from further analysis. Finally, 114 complete MADS-box genes were confirmed and named as *DlMADS1-DlMADS114* based on their genomic location in the chromosome (**Supplementary Table 2**). The length and molecular weight of 114 MADS-box proteins ranged from 64 AA and 7358.57 Da (*DlMADS85*) to 643 AA and 72378.92 Da (*DlMADS101*), with isoelectric points in the range of 5.06 to 10.49 (**Supplementary Table 2**). This result showed divergences in physicochemical properties among MADS-box family members in longan.

Phylogenetic and Evolutionary Analysis

To classify and examine the evolutionary relationship among MADS-box genes, we constructed a neighbor-joining (NJ) phylogenetic tree with alignments of longan and Arabidopsis MADS-box protein sequences. The result showed that *DlMADS* were divided into 5 groups, as in Arabidopsis (Parenicova et al., 2003; **Supplementary Figure 1**). Of these, 63 longan MADS-box genes were assigned to type I including 39 genes in *Mα*, 10 genes in *Mβ*, 14 genes in *Mγ*. 51 *DlMADS* were classified as type II including 36 $MIKC^C$ and 15 $MIKC^*$ type genes.

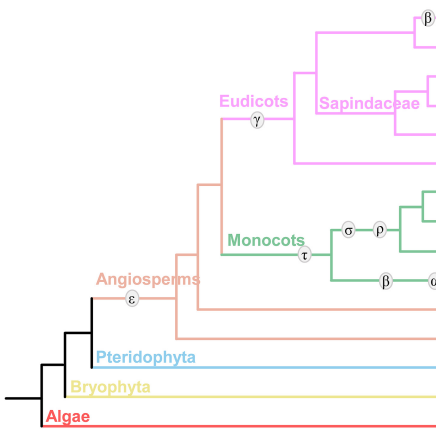
The MADS-box family is widely distributed in plants, and to understand the gene evolution in longan, we searched for MADS-box genes in the genomes of 15 representative plant species for comparative genomic analysis. Nine of them have been previously reported (see Methods), while MADS-box genes of the other six plant species were identified and classified from their genomes. As shown in **Figure 1**, both type I and type II have single genes in *C. reinhardtii* (Algae), supporting the view that these two types of MADS-box genes are very ancient and originated before the origin of Embryophyte. The MADS-box genes were expanded in Angiospermae due to the ε-whole-genome duplication (WGD). Compared to type I, the number of type II genes is relatively conserved in monocots and eudicots. In addition, longan has the largest number (114) of MADS-box genes, similar to litchi.

In order to further classify the $MIKC^C$ genes, two phylogenetic trees were constructed using MADS-box proteins from Arabidopsis, longan, litchi, and *Xanthoceras sorbifolium* (yellowhorn) using the maximum-likelihood (ML) and neighbor-joining (NJ) methods (**Figure 2** and **Supplementary Figure 2**). The structures of the two trees based on NJ and ML methods were similar, indicating a reliable subfamily division. The longan $MIKC^C$ genes could be divided into 12 subgroups with Arabidopsis genes as a reference, and no novel subfamily was found, indicating the conservation of the gene evolution in dicot species. The *Flowering Locus C* (*FLC*) clade of Arabidopsis contains one *FLC* gene and five homologs, *MADS-AFFECTING FLOWERING 1-5* (*MAF1-5*), which was attributed to the two rounds of Arabidopsis specific WGD. Only one gene was found in each Sapindaceae *FLC* clade, which was similar to the result in Citrus (Hou et al., 2013). Similar to *FLC*, subgroups *TT16* and *AGAMOUS-like 6* (*AGL6*) contained the minimum number (one) of longan type II genes. For the *SVP* clade, there are two genes (*SVP* and *AGL24*) in Arabidopsis, with the *SVP* subgroup consisting of the largest number (11) of $MIKC^C$ type genes in longan, indicating that some of the loss and duplication events probably occurred after the divergence of two the species. As *dormancy-associated MADS-box* (*DAM*) genes, which were highly homologous to *SVP* and *AGL24*, have been proven to affect dormancy in peach (Bielenberg et al., 2008), we collected reported *DAM* gene sequence data from *Prunus persica* (peach), *Prunus mume* (plum), *Pyrus pyrifolia* (pear), and *Malus domestica* (apple) to generate a phylogenetic tree with longan and litchi (**Supplementary Figure 3**). Except for the main *SVP* and *AGL24* clade, all the Rosaceae *DAM* genes cluster in one clade, with nine longan MADS genes (named as *DlSVP-LIKE*) and seven litchi MADS genes located in an individual cluster.

Genomic Distribution and Gene Duplication

The 114 MADS-box genes were unevenly distributed on 15 chromosomes of the longan genome (**Figure 3A**). At least 3 *DlMADS* were found in each chromosome, chromosome 11 had the largest number of family members at 20, and only

[13]https://sg.idtdna.com/pages/tools/primerquest



	Total	Type I	*Mα*	*Mβ*	*Mγ*	Type II	*MIKC*^{C}	*MIKC**
A. thaliana	107	61	24	21	16	46	39	7
C. clementina	84	35	24	6	5	49	44	5
D. Longan	114	63	39	10	14	51	36	15
L. chinesis	110	59	34	13	12	51	36	15
X. sorbifolium	79	32	19	6	7	47	32	15
V. vinifera	54	10	3	6	1	44	39	5
S. bicolor	65	30	26	2	2	35	33	2
Z. mays	75	32	27	3	2	43	39	4
O. sativa	75	32	13	9	10	43	38	5
M. acuminata	96	31	-	-	8	65	60	5
N. colorata	70	41	32	5	4	29	25	4
A. trichopoda	34	12	6	4	2	22	20	2
S. moellendorffii	18	8	-	2	-	10	7	3
P .patens	24	7	-	-	-	17	6	11
C. Reinhardtii	2	1	-	-	-	1	-	-

Whole Genome Duplication

FIGURE 1 | The MADS-box gene family in 15 species. The evolutionary relationship of 15 species on the left of the figure was obtained from NCBI (https://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi). The right of the figure shows the number detail of the MADS-box family in each species. Type I MADS-box genes contain three groups: *Mα*, *Mβ*, *Mγ*, and type II genes are subdivided into *MIKC*^{C} and *MIKC** groups.

three genes were located on chromosomes 4 and 12. Gene duplication has been considered the driving force for species evolution, and WGD events have probably occurred in many eukaryotes, sometimes more than once (Van de Peer, 2004). Previous research has established that over 90% of the increase in the number of regulatory genes was caused by the three WGD events in Arabidopsis (Maere et al., 2005). In this study, we investigated gene duplication events in Arabidopsis, longan, litchi, and yellowhorn genome using MCScan X. Of the genes of the longan MADS-box family, 44 (38.6%) originated from WGD or segmental duplication, 20 (17.5%) appeared to have been created through tandem duplication, 22 (19.3%) were proximal duplicated genes and 28 (24.6%) were dispersed duplicated genes (**Table 1**). In Arabidopsis, nearly half of the MADS-box genes (46.7%) originated from dispersed duplication, while tandem duplication events were more widespread in the MADS-box genes which indicates they made a valuable contribution to the evolution of the longan MADS-box family. Of significance, at least 26 genes of the *Mα* group were clustered on special regions of chromosomes 2, 7, and 13, indicating that tandem duplications were the main force driving the expansion of this MADS-box group.

In Sapindaceae, type II gene numbers were generally conserved, while type I gene numbers varied. Longan and litchi have consistent numbers of MADS-box genes in all the subgroups with the exception of *Mβ* and *Mγ* (**Figure 1**). Yellowhorn had much fewer type I MADS-box genes (32) than both longan (63) and litchi (59). Therefore, we constructed the comparative syntenic maps among longan, litchi, and yellowhorn (**Figure 3B**). For type I, 44 orthologous gene pairs were identified between longan and litchi, indicating close relationships. In comparison to longan and litchi genomes, gene concentration was observed in which was mainly caused by the loss of type I genes in the corresponding longan chromosomes 1, 2, 3, 6, 11, 13, 14, and 15. Moreover, the synonymous substitution rates (Ka/Ks) of the gene pairs were calculated to identify the evolutionary forces. All of the 58 orthologous gene pairs had Ka/Ks < 1 (**Supplementary Table 3**), suggesting that purifying selection may be the dominant force driving the evolution of Sapindaceae type I MADS-box genes.

Gene Structure and Conserved Motifs of Longan MADS-Box Family

To assess the structural diversity of longan MADS-box genes, the GSDS program was used to display intron-exon organization. The results showed that the structures of type I genes were short and simple with the presence of 0–2 introns except for *DlMASD101* with 9 introns (**Supplementary Figure 4**). The full-genome lengths of *DlMADS40* and *DlMADS42* were significantly longer at 6503 and 9040 bp. Through the survey of transposable element (TE) sequences, we found that 9 and 5 TEs were located in *DlMADS40* and *DlMADS42*, respectively (**Supplementary Table 4**). Compared with type I, type II genes contain more exons in the range of 0–12, which is in contrast to the previous report that *MIKC* genes have a common structure of 1–6 exons (Johansen et al., 2002). In addition, the average length of *MIKC** genes (3,900 bp) was shorter than *MIKC*^{C} genes (8,800 bp).

Since *MIKC* type genes with complex structures play functional roles in developmental processes in plants (Becker and Theissen, 2003), we analyzed the conserved motifs of longan *MIKC* type proteins using the MEME program. A total of 10 conserved motifs were identified among 51 *MIKC* protein sequences (**Figure 4** and **Supplementary Table 5**) and the number of motifs in *DlMADS* ranged from 1 to 7 and each subfamily had similar motif compositions. Motif 1 which was annotated as MADS domain was found in nearly all protein sequences except for *DlMADS86*, *84*, *87* in the *SVP* class and

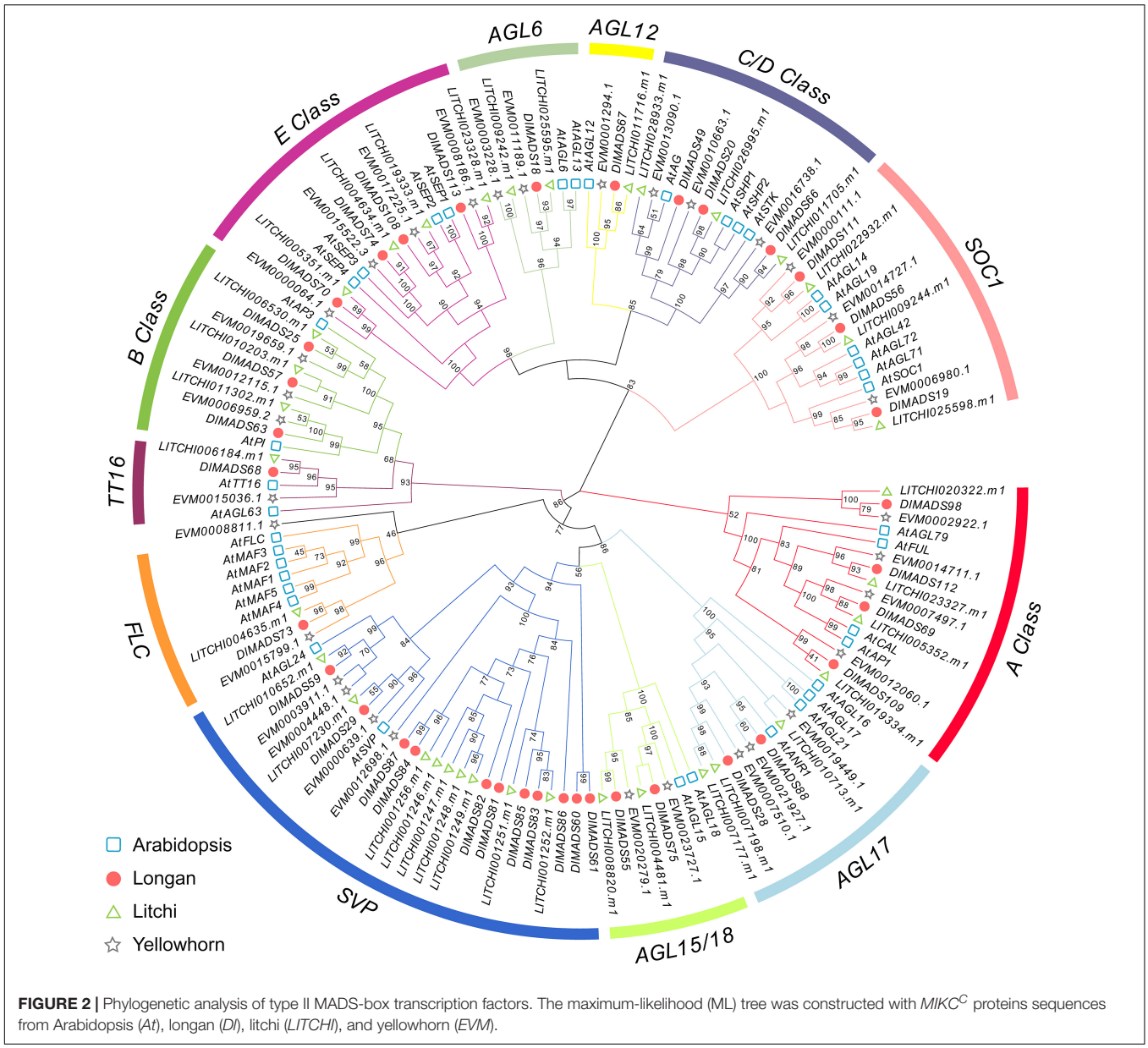


FIGURE 2 | Phylogenetic analysis of type II MADS-box transcription factors. The maximum-likelihood (ML) tree was constructed with *MIKC*^{C} proteins sequences from Arabidopsis (*At*), longan (*Dl*), litchi (*LITCHI*), and yellowhorn (*EVM*).

DlMADS1 in the *MIKC** class, with the distribution of the unknown motif 4 being similar to motif 1, suggesting that these two motifs play crucial functions. Motif 2, 3, and 7 may be fragments of the K-box and are distributed in most *MIKC*^{C} members, but motif 5 annotated as K-box was only found in *MIKC** class. The result reveals that the primary differences between *MIKC*^{C} and *MIKC** proteins are variations in the K domain.

Cis-Elements and Potential Transcription Factor Binding Sites

Transcription factors (TFs) control and regulate gene expression through binding cis-elements in the promoters of target genes. To investigate the regulatory gene networks of the MADS-box family, we analyzed the cis-elements of the upstream 1,500-bp sequences of *DlMADS* based on the PlantCARE database. Cis-elements associated with 11 biological processes such as light-responsiveness and the circadian clock were annotated in *DlMADS* (**Supplementary Table 6**). Of note, the number of functional elements in *MIKC*^{C} type genes is significantly less than that of other genes (**Figure 5A**). Light-responsive boxes existed in all MADS-box genes. Hormone-related (ABRE, TGA-element, AuxRR-core, GARE-motif, P-box, etc.) and defense and stress-responsive (WUN-motif, LTR, TC-rich repeats, ARE) cis-elements were found in most family members. Additionally, *DlMADS1* belonging to the *MIKC** group had the most cis-elements amongst the gene families. In the *MIKC*^{C} subfamily, the hormone-related elements were more frequently located in *SEP*, *FLC*, and *SVP* clades.

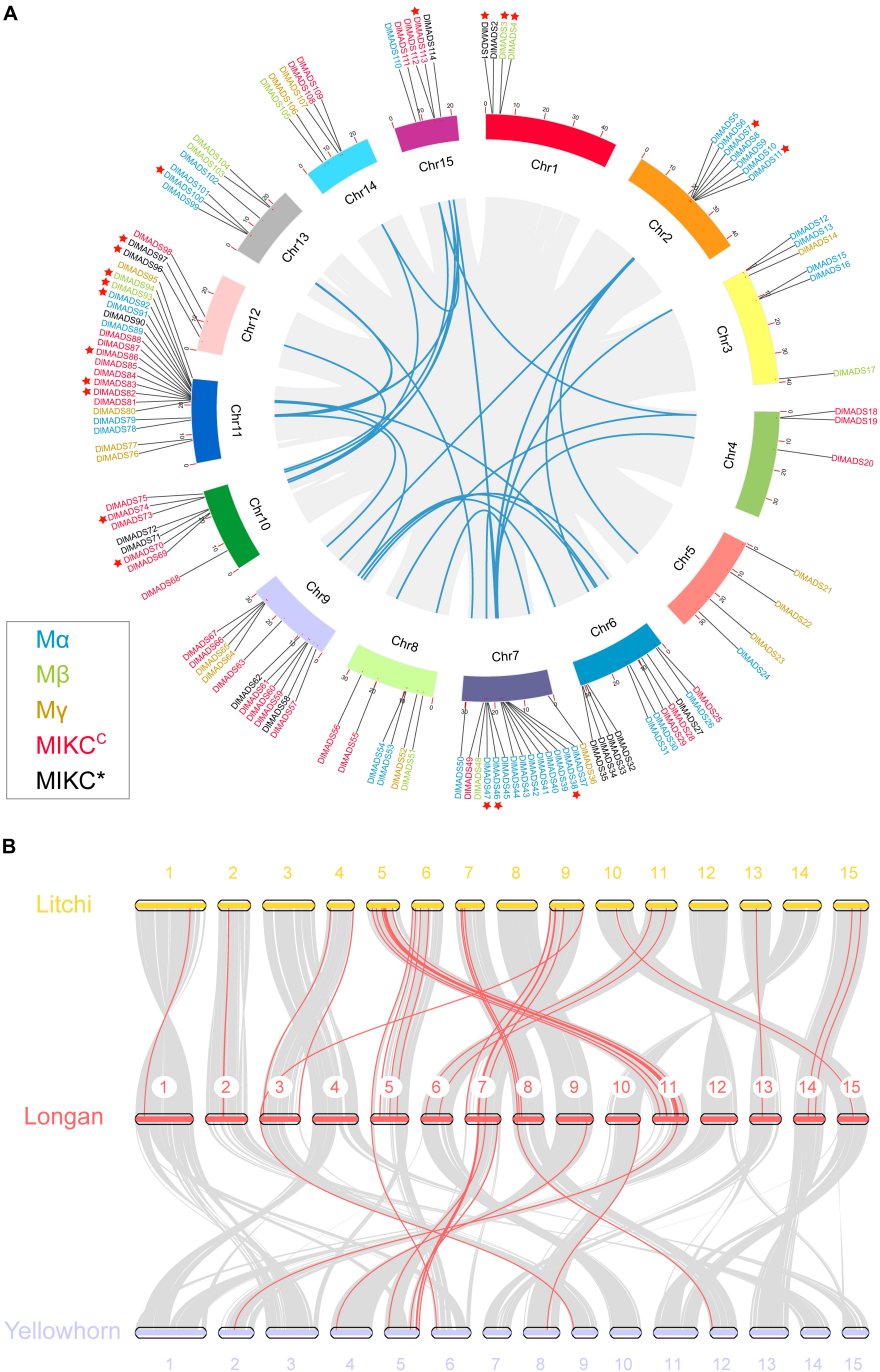


FIGURE 3 | Chromosomal location and synteny analysis of the MADS-box genes. **(A)** A total of 114 MADS-box genes are located in 15 chromosomes with different colors. Gene pairs of WGD or segmental duplication are linked using blue lines. Tandem duplication genes are marked by red stars. Gene labels of five types: *Mα*, *Mβ*, *Mγ*, *MIKC**, and *MIKC*^{C} are denoted by blue, green, yellow, red, and black, respectively. **(B)** Synteny relationship of type I MADS-box genes from longan, litchi, and yellowhorn genomes. Gray lines in the background indicate the collinear blocks within two genomes. Red lines highlight the syntenic type I MADS-box gene pairs.

Transcription factors can directly bind to specific sequences in promoters of target genes called binding sites to affect gene expression. In order to further identify the transcription factors which potentially control the *MIKC*^{C} type genes, the promoters of 36 *MIKC*^{C} genes were processed using MEME software to locate cis-motifs. As a result, 20 enriched motifs were found and a total of 26 TFs along with their DNA bind sites were identified in 7 motifs by comparison

TABLE 1 | Numbers of MADS-box genes from different origins in Arabidopsis, longan, litchi, and yellowhorn genomes.

	Number of MADS-box genes	Number of genes from different origins (percentage)				
		Singletons	WGD/Segmental	Tandem	Proximal	Dispersed
Arabidopsis thaliana	107	4 (3.7%)	24 (22.4%)	11 (10.3%)	18 (16.8%)	50 (46.7%)
Dimocarpus longan	114	0 (0.0%)	44 (38.6%)	20 (17.5%)	22 (19.3%)	28 (24.6%)
Litchi chinensis	114	1 (0.1%)	18 (15.8%)	27 (23.7%)	14 (12.3%)	54 (47.4%)
Xanthoceras sorbifolium	79	1 (0.1%)	22 (27.8%)	13 (16.5%)	8 (10.1%)	35 (44.3%)

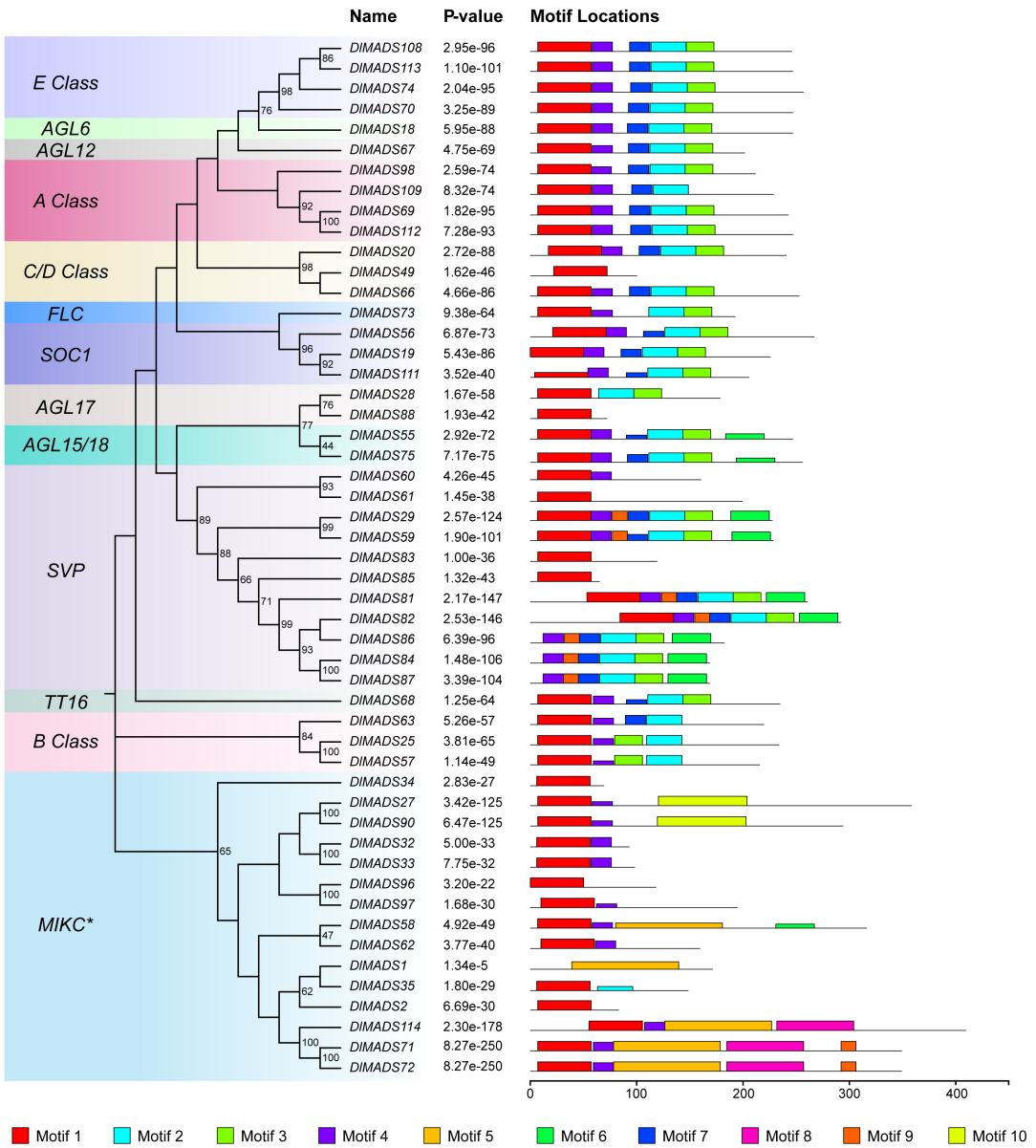


FIGURE 4 | Conserved motif compositions of longan type II MADS-box proteins. The neighbor-joining tree was constructed with the aligned protein sequences of longan type II MADS-box genes. 10 motifs were identified and displayed in different colors.

with the JASPAR database (**Figure 5B** and **Supplementary Table 7**). 21 TFs were *C2H2* zinc finger proteins particularly the DNA-binding *One Zinc Finger* (*Dof*) transcription factor family. For example, *CYCLING DOF FACTOR 5* (*CDF5*) and *COGWHEEL1* (*COG1*) binding sites were found in 4 motifs. Other TFs belong to *BARLEY B RECOMBINANT/BASIC*

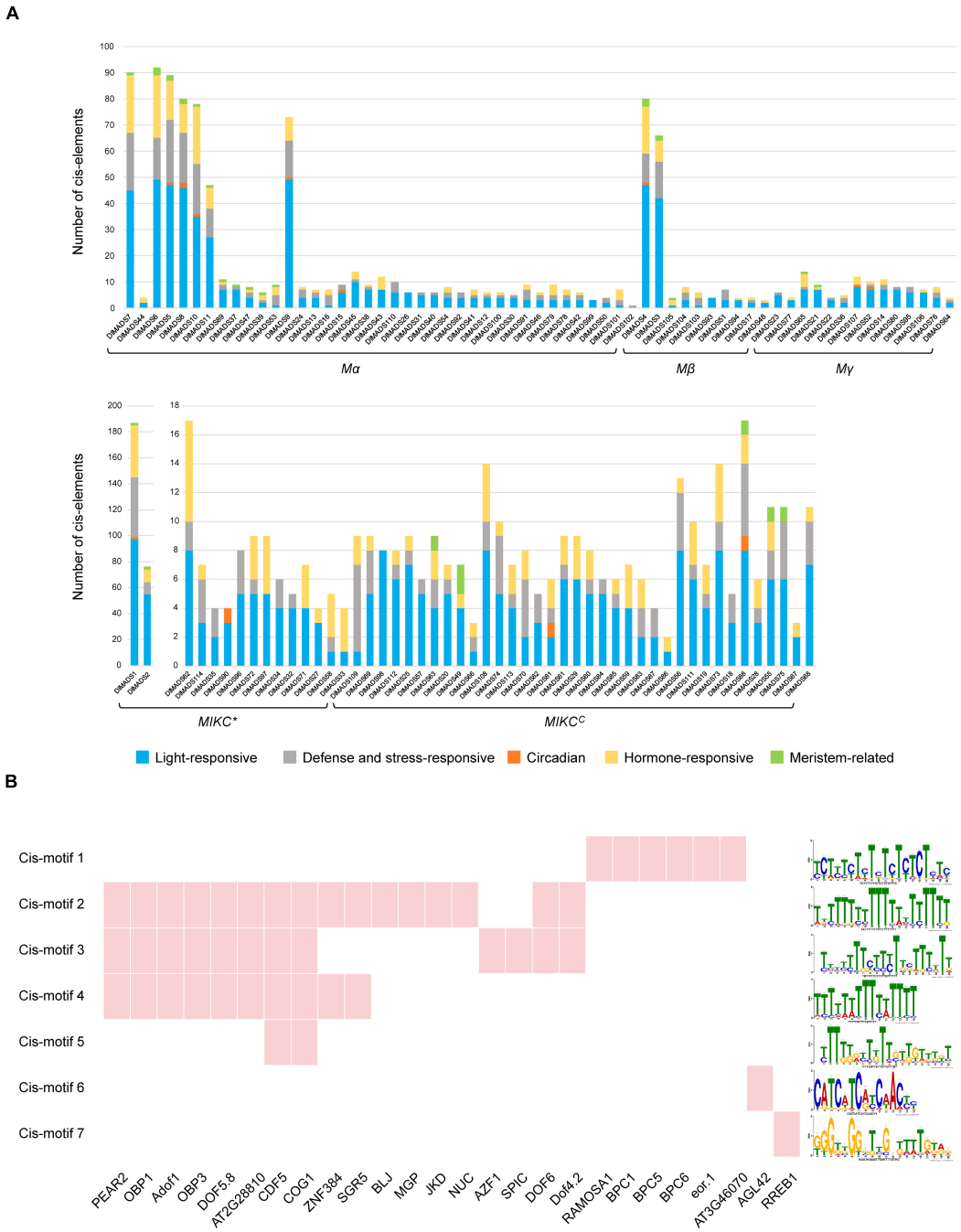


FIGURE 5 | Prediction of cis-elements and transcription factors of *DlMADS*. **(A)** The number of cis-elements related to light-responsive, defense and stress-responsive, circadian, hormone-responses and meristem. The upstream 1.5 kb sequences of all longan MADS-box genes were analyzed through PlantCARE. **(B)** The cis-motifs and transcription factors potentially target 36 *MIKC*^{C} genes. Red indicates the transcription factor binding sites were found in promoter motifs of *MIKC*^{C} genes.

PENTACYSTEINE (*BBR/BPC*), MADS-box family and, ETS-related genes.

A Proposed ABC(D)E Model in Longan

Extensive research showed that five classes of homeotic genes called the ABCDE model, control floral development at the molecular level. Phylogenetics and classification based on the model plant as reference allowed the identification of the putative functions of longan MADS-box genes. As shown in **Figure 6A**, only one copy of an A-class functional MADS-box gene was found in Amborella and four homologs were confirmed in rice, Arabidopsis, and the three Sapindaceae species. Phylogeny

indicated that A-class genes were present in a one-to-one orthologous pattern in the Sapindaceae species. *AP1* and *CAL* originated from a recent duplication event and have both partially redundant and unique functions (Alvarez-Buylla et al., 2000), and only one gene of longan (*DlMADS69*) was found in *AP1/CAL* clade, named as *DlAP1*. In addition, the number of *AP2* genes in the *AP2/EREBP* family was highly conserved in tested Angiosperm (**Figure 6A**). Both *AP3* and *PI* which diverged in the ancestors of angiosperms experienced loss events in eudicots, and C and D class genes formed a very close sister group, with each Sapindaceae species having one *AG*, one *STK,* and one *SHP* gene. *SEP* genes have roles throughout five whorls of floral organs and were duplicated during angiosperm evolution. Four *SEP* homologs were found in longan, litchi, and yellowhorn. Overall, these results demonstrated that the number of ABCDE genes in Sapindaceae was similar to the model plant Arabidopsis.

In order to investigate the correlation between ABCDE genes and longan floral organs, we analyzed the expression patterns of certain genes in mature four whorls of floral organs (sepals, petals, stamens, carpel) and ovary by performing qRT-PCR (**Figures 6B,C**). Although twelve selected genes could be detected in most organs, they showed high specificity in different organs. For A-class, *DlAP1* was mainly expressed in the first three whorls and *DlMADS112* (*DlFUL)* was mainly expressed in carpel and ovary. *DlMADS63* (*DlPI*) of B class was only detected in petal and stamen, *DlMADS25* (*DlAP3-1*) also showed relatively high expression in these two organs, even though it could be found in all five organs. The only C class gene *DlMADS49* (*DlAG*) was detected in sepal, stamen, carpel, and ovary, especially in stamen and carpel. Two D-class genes, *DlMADS20* (*DlSHP*) and *DlMADS66* (*DlSTK*) were expressed in reproductive organs (stamen and pistil) and showed relatively high expression levels in the ovary which contains ovules. Interestingly, four members of the E class showed different expression patterns in floral organs suggestive of diversified gene function. The expression levels of *DlMADS74* and *113* (*DlSEP1* and *3*) were well-proportioned in five orangs. However, *DlMADS108* (*DlSEP2*) showed a similar expression pattern with *DlAP1*, and *DlMADS70* (*DlSEP4*) was preferentially expressed in the ovary like *DlSTK*.

Besides the ABCDE genes, *AGL6* has been demonstrated to control floral development. In Arabidopsis flowers, *AGL6* could be detected in all floral organs and in developing ovules (Koo et al., 2010). However, the transcripts of *DlMADS18* (*DlAGL6)* were only detected in the first three whorls of organs (sepals, petals, and stamens), which suggested that this gene may not control longan ovule development. These results allowed us to predict the ABC(D)E model in longan (**Figure 6D**).

Tissue-Specific Expression Patterns of *DlMADS*

To investigate the expression pattern of longan MASD-box genes, previously published transcriptome data from multiple tissues including root, stem, leaf, dormant bud (before the emergence of floral primordia), flower bud, flower, young fruit, pericarp, pulp, and the seed of "SJ" longan were used for analysis (Lin et al., 2017; Jue et al., 2019). In line with previous studies (Kofuji et al., 2003; Duan et al., 2015), almost all genes from type I displayed low expression levels (FPKM value < 1) in nine tissues (**Supplementary Table 8**). In addition, *DlMADS21* was highly expressed specifically in roots and *DlMADS53* showed extremely high expression in pulp (**Supplementary Figure 5**), which suggests that these two genes play roles in root and fruit development, respectively. Moreover, *DlMADS45* was primarily expressed in root, stem, and seeds. *DlMADS65* was lowly expressed in multiple tissues. In the *MIKC** group, *DlMADS71*, *72*, *114* with more protein motifs in the K-box region than others showed higher expression in nearly all tissues, which provides evidence to the fundamental role of the K-box domain. The expression level of *DlMADS27* and *DlMADS58* in flower buds and flowers were higher than other organs, indicating crucial roles in the flowering process.

MIKC^C type genes have important functions in the productive organs in plants (Becker and Theissen, 2003). In longan, genes of *MIKC^C* displayed a wide range of expression levels in different tissues, among which 11 (30.6%), 18 (50.0%), 17 (47.2%), 17 (47.2%), 24 (66.7%), 20 (55.6%), 16 (44.5%), 16 (44.5%), 8 (22.2%), 14 (38.9%), and 16 (44.5%) were expressed (FPKM value > 1) in root, stem, leaf, dormant bud, flower bud, flower, young fruit, pulp, pericarp and seed, respectively (**Supplementary Figure 6**). Although expression patterns of *DlMADS* were conserved within each subfamily, several genes showed different expression levels in tested tissues such as *DlAP1*, *DlFUL,* and *DlMADS109* in the A-class group. Our analysis showed that expression patterns of *MIKC^C* genes in longan tissues were classified into four clusters (**Figure 7A**). In cluster 1, most genes show little expression but several genes such as *SVP* class members were expressed in leaf, stem, and dormant bud. *DlMADS67* and *DlMADS88* were specifically expressed in roots. In cluster 2, *DlMADS* was mainly expressed in vegetative tissues. *DlMADS73* (*DlFLC*), *DlFUL,* and *DlMADS29* (*DlSVP*) of cluster 3 displayed extensive-expression levels in nearly all tested tissues. In cluster 4, *DlAGL6* and 9 genes of the ABCDE classes were mainly expressed in the reproductive organs including flower bud, flower, fruit, pericarp, and seed. From this finding, it is clear that *MIKC^C* type *DlMADS* might perform significant roles in flowering and fruit development as opposed to other organs.

To further survey the expression patterns of key *DlMADS* having reliable transcriptional support, quantitative RT-PCR was performed using total RNA isolated from leaf, dormant bud (before the emergence of floral primordia), floral bud (floral primordia), flower, and pulp (**Figure 7B**). Among seven selected genes, *DlAP1*, *DlMADS57* (*DlAP3-2*), and *DlAGL6* were only detected in floral or flower tissues, suggesting a unique function in flowering. *DlSEP3* showed high expression in the floral bud, flower, and pulp tissues. Moreover, *DlMADS53* showed specific high expression in pulp, consistent with the result of RNA-seq.

Differential Expression of *DlMADS* in Off-Season Flower Induction

The compound potassium chlorate ($KClO_3$) has been widely used to induce flowering in longan since the last century (Subhadrabandhu and Yapwattanaphun, 2001). In this study, we

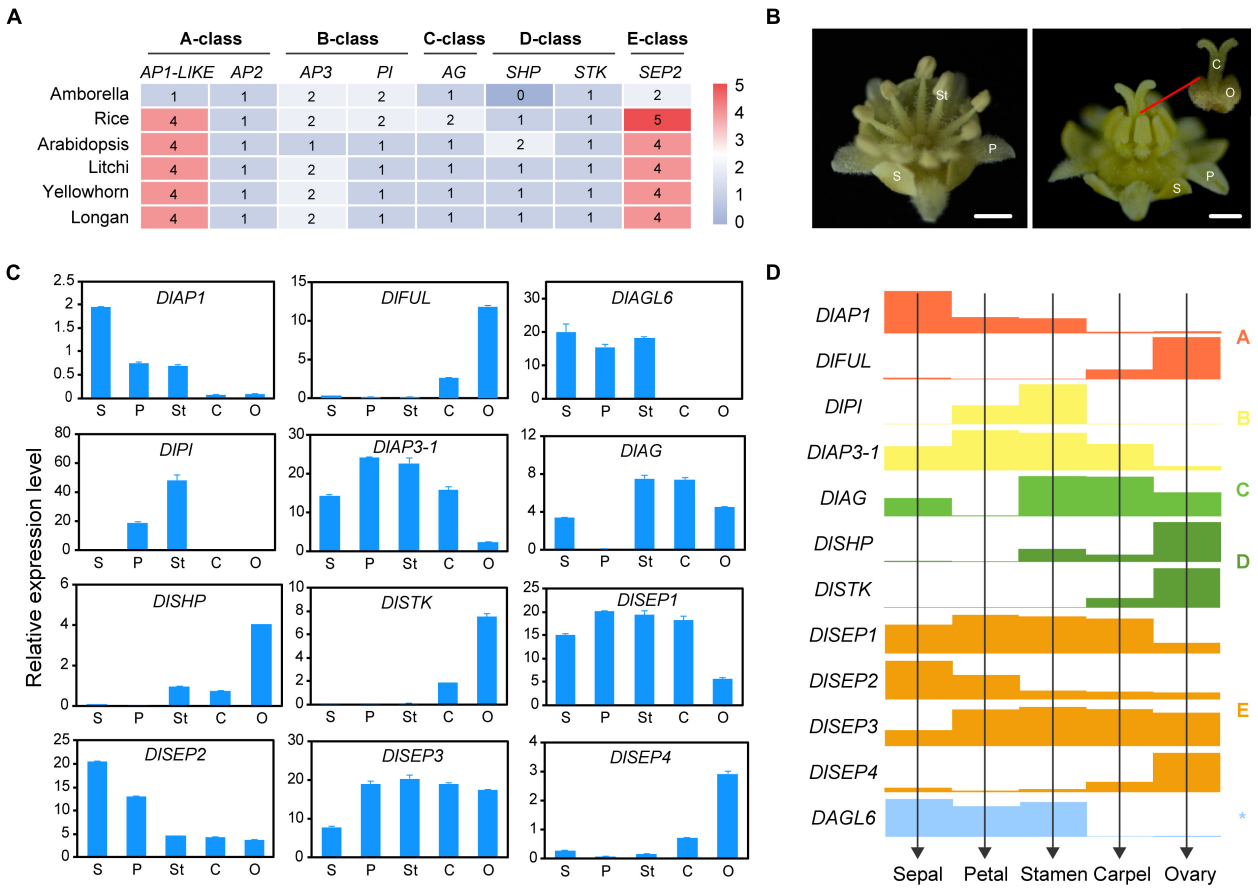


FIGURE 6 | The deduced ABCDE model in longan. **(A)** The heat map shows the number of ABCDE classes genes in Amborella, Arabidopsis, rice, and longan. **(B)** Mature male (left) and female (right) flowers of longan. Five types of floral organs including sepal (S), petal (P), stamen (St), carpel (C), ovary (O) were collected for gene expression analysis. **(C)** Expression analysis of ABCDE classes genes in floral organs in **(B)** by using qRT-PCR. **(D)** The ABCDE model in longan is proposed based on the gene expression levels (bar heights) from **(C)**.

detected the expression of MADS-box genes in apical buds at ten-time points using untreated controls (CK) and $KClO_3$ treatments (see Methods). A total of 34 genes were expressed in more than one stage (FPKM > 1), including 26 *MIKC*^{C} type genes, 3 *MIKC** genes, and 5 genes of type I (**Supplementary Table 9**).

Among these, we found 14 genes showed significant differential expression levels (| logFC| > 2 and *P*-value < 0.05) after $KClO_3$ treatment, including the ABCE model genes and other flowering time integrators (**Figure 8A**). *DlAP1*, *DlMADS98,* and *DlSEP4* were up-regulated from day 41 to day 54, especially *DlAP1* with the highest FPKM value. It is interesting to note that *DlAP3-1*, *DlAP3-2*, *DlPI*, *DlAG*, *DlSEP1/2/3*, and *DlAGL6* showed the same patterns with upregulation on day 54. Therefore, these up-regulated genes could be considered as promoting factors for off-season flowering induction. In addition, *DlMADS111*, *DlFLC*, and *DlMADS82* displayed a decreased pattern from day 35 to day 54, suggesting that these may inhibit longan flowering in this time period.

To assess the relationships between *DlMADS* and other genes throughout the genome that may affect off-season flowering, the expressed genes were classified into 16 modules by weighted gene co-expression network analysis (WGCNA) (**Supplementary Figure 7**). The result showed that 14 differentially expressed MADS-box genes were distributed in four modules (MEyellow, MEbrown, MEturquoise, and MEgreenyellow). Based on the differentially expressed MADS-box genes, we constructed the co-expression networks and found that 113 genes were presumed to interact with them (**Figure 8B**). Based on the *Arabidopsis thaliana* flowering interactive database[14], we found five flower-related genes interacted with MADS-box genes in the networks including *UNUSUAL FLORAL ORGANS* (*UFO*), *TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTOR* (*TCP*), *CRABS*, *FLOWERING PROMOTING FACTOR1* (*FPF1*), and *HOTHEAD* (*HTH*) (Kania et al., 1997; Krolikowski et al., 2003; Hepworth et al., 2006; Gross et al., 2018; Zhao et al., 2020). Besides these flower-related genes, (*UDP-Glycosyltransferase*-2) *UGT-2*, (*SPATULA*) *SPT*, and *SWEET1* were predicted to co-expressed with the MADS-box genes although have not yet been reported for the function related to flower. Whether these genes play functions in longan flower induction and development

[14]http://www.phytosystems.ulg.ac.be/florid/

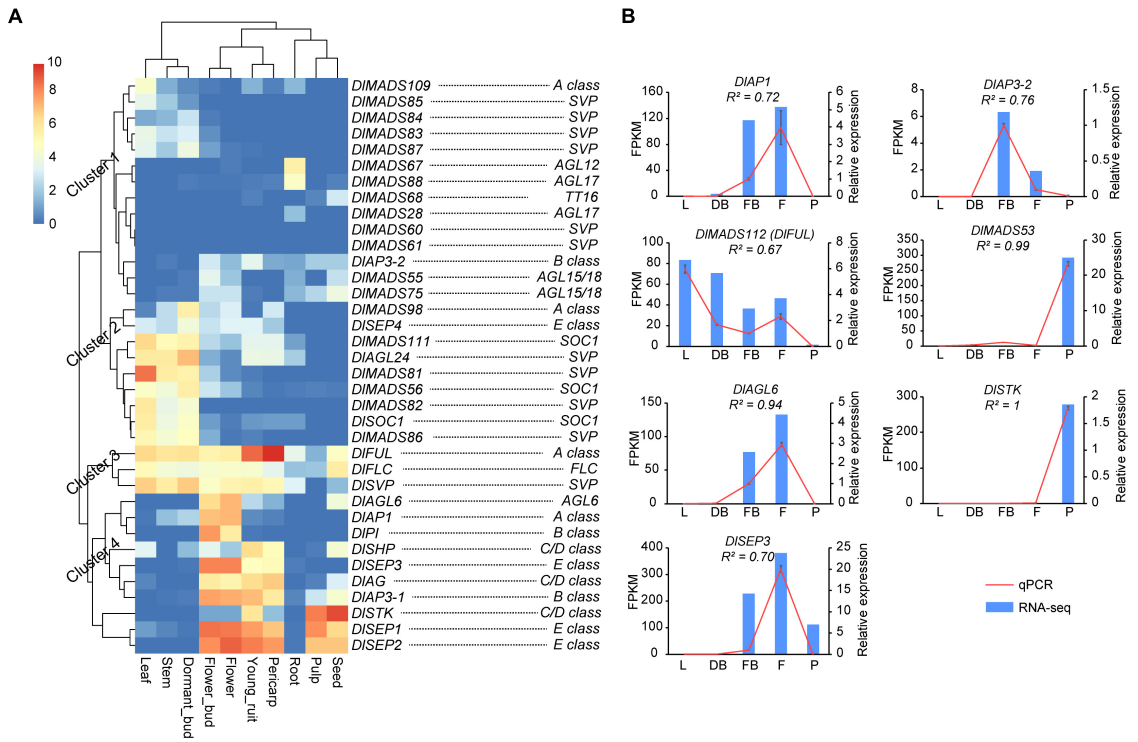


FIGURE 7 | The expressional pattern of *MIKC*^{C} genes in longan. **(A)** Expression heat maps of *MIKC*^{C} genes in nine tissues (root, stem, leaf, dormant bud, flower bud, flower, fruit, pericarp, and seed) of "SJ" longan Expression values were normalized by $\log_2$(FPKM+1). **(B)** Detection of MADS-box genes in longan based on qRT-PCR. Six MADS-box genes were selected to detect the expression level in leaf (L), dormant bud (DB), floral bud (FB), flower (F), and pulp (F) of "XC" longan. Data were normalized to Actin1 and Actin2 genes, and the vertical bars indicate standard deviation.

needs more experiments to confirm. In addition, 8 *DlMADS* (*DlMADS98*, *DlAG*, *DlSEP1/2/3*, *DlPI*, *DlAP3-2*, and *DlAGL6*) were found in the green-yellow module and *DlAGL6* was the top hub gene with the highest kME value (0.99) indicating the central role in flower development (**Figure 8B**).

Moreover, we also compared the expression levels of *DlMADS* between "SX" and "SJ" during three floral transition stages (T1: dormant bud; T2: floral primordia; T3: floral organ formation). As a result, seven MADS-box genes were significantly differentially expressed in two accessions (**Supplementary Figure 8**). The expression of *DlMADS111*, *DlAGL24*, *DlFLC*, and *DlSOC1* was higher in "SX_T1" than that in "SJ_T1," and they were all downregulated in the following two stages except for *DlSOC1*. *DlAP1* was upregulated in T1 and T2 in both accessions but showed higher expression levels in "SX" than that in "SJ." On the contrary, the *DlMADS98* displayed a higher expression in "SJ" than that in "SX" in the entire stages.

Expression Profiling of *DlMADS* During Fruit Development

It was reported that the MADS-box gene family have functions in fruit development and ripening (Moore et al., 2002), and our present study analyzed the expression pattern of *DlMADS* in six fruit developmental stages at 80, 100, 110, 120, 130, and 140 DAF (days after flowering of female flower) of 10 years "XC" tree by RNA-seq. 21 MADS-box genes were expressed in at least one stage of fruit development (**Figure 9A** and **Supplementary Table 10**). Ten genes (*DlMADS75*, *70*, *59*, *55*, *72*, *18*, *114*, *29*, *49*, *68*) were only expressed at low levels in the first two stages of fruit development (DAF 80 and DAF 100). *DlAP3-1*, *DlAP3-2*, and *DlMADS71* were down-regulated gradually during developmental stages and *DlFLC* showed the reverse trend. *DlSTK* and *DlSHP* exhibited the highest expression levels in DAF 80, while the expression of *DlSTK* rose in later stages and *DlSHP* decreased until it was not expressed at all. Increasing expression from DAF 80 to DAF 130 was found in *DlSEP1* and *DlSEP2* indicating that they may be involved in the fruit ripening process. It is significant that *DlMADS53* (*Mα*) showed an extremely high expression level suggesting important functions in longan fruit development. The RNA-seq results of *DlAP3-2*, *DlSTK*, *DlFLC*, *DlFUL*, and *DlMADS53* were corroborated by qRT-PCR. The relative expression levels were in agreement with the FPKM, confirming the accuracy of our transcriptomic analysis ($R^2 > 0.9$; **Figure 8B**).

DISCUSSION

Given the developments in genome sequencing, there are numerous reports on the important MADS-box family in various plants such as Arabidopsis (Parenicova et al., 2003),

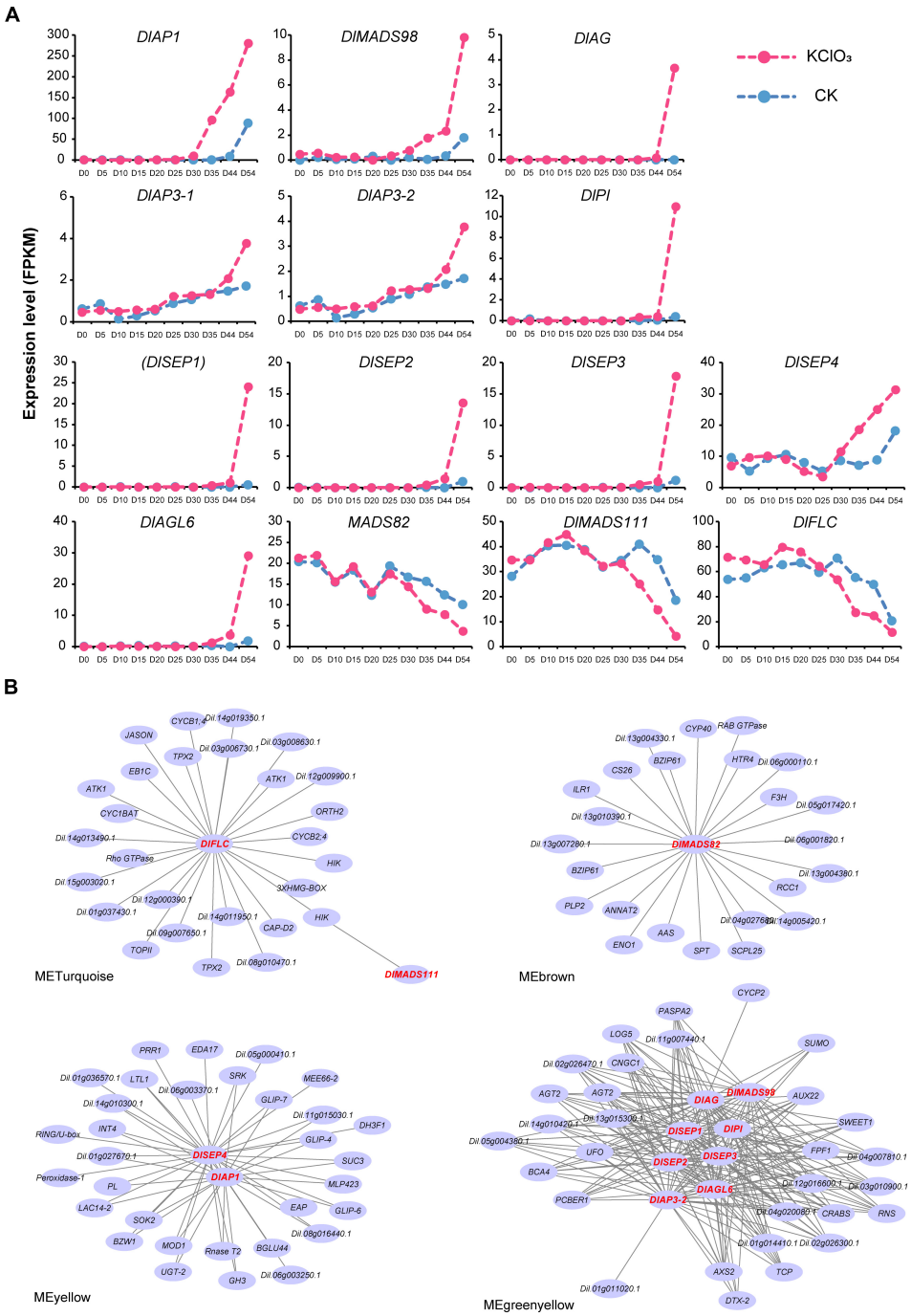


FIGURE 8 | Transcript analysis of longan MADS-box genes in off-season flower induction. **(A)** Expression profile of differentially expressed MADS-box genes during 10 stages (day 0~54) under CK and $KClO_3$ treatment. **(B)** Co-expression networks of differential expression MADS-box genes during off-season flower induction. Each node represents a gene and the lines between nodes represent co-expression correlations.

rice (Arora et al., 2007), soybean (Fan et al., 2013), and banana (Liu et al., 2017). In this study, we completed the identification of the MADS-box gene in "SX" longan genome for the first time, and 114 MADS-box members were identified and classified into type I: *Mα* (39), *Mβ* (10), *Mγ* (14), and type II: *MIKC*^{C} (36), *MIKC** (14) (**Supplementary Table 2**). Combined with the relative representative plant species, it is possible to understand the evolution of MADS-box genes.

Evolution of MADS-Box Family

In this study, we identified the MADS-box genes from 15 representative species (**Figure 1**). The primitive lower plant,

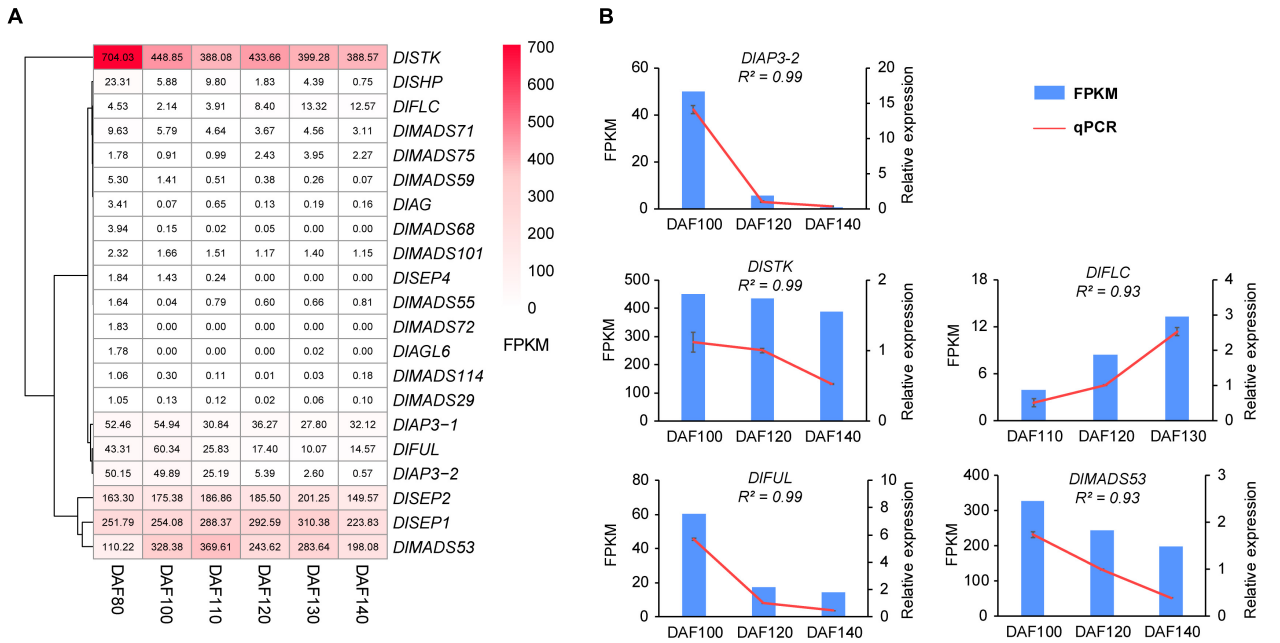


FIGURE 9 | The expression pattern of differentially expressed *DIMADS* during the fruit development stage. **(A)** Heat maps based on FPKM for differentially expressed *DIMADS* at three fruit developmental stages (80∼140 days after flowering). **(B)** Detection of five differentially expressed *DIMADS* at three selected stages by using qRT-PCR. The bar and line graphs are derived from RNA-seq and qRT-PCR data, respectively.

Chlamydomonas reinhardtii, only contained two genes belonging to both type I and type II, in accordance with the view that these two types of genes appeared by a duplication event before the divergence of plants and animals (Alvarez-Buylla et al., 2000). Overall, the number of MADS-box genes in Angiosperms is larger than that in Pteridophyta, Bryophyta, and Alga. A large expansion of $MIKC^C$ genes was found in basal angiosperms, given that flowering is a typical feature of high plants, we predict $MIKC^C$ genes may be endowed with new functions as plants evolved especially in flowering and development. In monocots and eudicots, type II genes showed relative conservation and the number of type I genes varied greatly in eudicots. It was found that longan contained the largest number of family genes, and we speculate that members of the MADS-box family were likely more active for tandem duplication after the γ-WGD event in longan than other eudicot species (**Table 1**). Furthermore, the degree of retention of type I genes seems to be higher than in other angiosperms. $MIKC^*$ genes expanded much more and they mainly originated from WGD/Segmental and dispersed duplication.

In Sapindaceae, the three examined genomes have no additional WGD after the γ-WGD, while, both longan and lichee have large numbers of MADS-box genes as the tandem duplication events frequently occurred in the longan genome. This phenomenon indicated that longan and lichee probably have complex molecular mechanisms for flower development due to the interaction of large sets of MADS-box genes for the regulation of gene networks. Sapindaceae could be classified into two subfamilies including Sapindoideae and Dodonaeoideae (Harrington et al., 2005). The divergence of the two subfamilies resulted in a great variation in the number of the type I genes particularly for *M*β and *M*γ (**Figure 1**), which was probably caused by the functional redundancy of the two subgroups of MADS-box genes since these genes have few reports detailing their functions (Masiero et al., 2011). $MIKC^C$ genes were demonstrated to be involved in floral organ identity, the control of flowering time, and seed development. The relatively consistent number of $MIKC^C$ genes in Sapindaceae is likely attributed to the functional conservation of these genes in floral organ development.

Dof Family May Regulate $MIKC^C$ Genes in the Flowering Process

The potential transcription factors which control $MIKC^C$ genes were identified, and 26 TF binding sites were found in promoters of most $MIKC^C$ genes (**Supplementary Table 7**). The majority of TFs contain *C2H2* zinc finger domain such as the *Dof* family, which is involved in photoperiod inducement of flowering and flowering development (Yanagisawa, 2002). In Arabidopsis, *CYCLING DOF FACTOR* (*CDF*) is the well-known inhibitory factor of flowering which can repress the expression of *FT* and *CONSTANS* (*CO*), it can be degraded by ubiquitin-protein *KELCH REPEAT, F-BOX 1* (*FKF1*) (Imaizumi et al., 2005). *OBF4 Binding Protein 3* (*OBP3*) and *Cogwheel 1* (*COG1*) modulate phytochrome to further influence flowering (Park et al., 2003; Ward et al., 2005). Other research showed that maize transcription factor *Zmdof1* regulates pollen formation and development (Chen et al., 2012). In rice, overexpressing *OsDof12* led to early flowering and up regulation of *OsMADS14*

(Li et al., 2009). The previous study has demonstrated that *Dof* proteins were able to recognize (A/T)AAAG sequence as the core motif (Chen et al., 1996; Yanagisawa and Schmidt, 1999). In this study, similar binding sites of *CDF5*, *COG1*, *FOXP1*, *OBP3,* and other *Dof* genes were found in cis-motif 2, 3, 4, and 5 (**Figure 5B**). It could be speculated that *Dof* TFs may directly regulate *MIKC^C* genes in longan, however, this speculation needs further investigation.

BBR-BASIC PENTACYSTEINE (*BPC*) identified in Arabidopsis is a small transcription factor family of seven members (Meister et al., 2004). *BPC1* binds to the *STK* promoter at GA consensus sequences and regulates the expression of *STK* (Kooiker et al., 2005). In addition, chromatin immunoprecipitation analysis showed that *BPCs* also bind to the GA boxes *in vivo*, and their mutation can result in suppressed expression of the *STK* promoter (Simonini et al., 2012). Our data showed that three genes of the *BBR-BPC* family probably bind sequences in cis-motif 1, which provides evidence that *BPC* binding sites are important for *STK* and other *MIKC^C* genes (**Figure 5B**). Moreover, *AGAMOUS-LIKE 42* (*AGL42*) is a member of the *SOC1* subfamily and controls the floral transition in the axillary meristem, and *SOC1* directly binds to CArG-boxes from other *SOC1* class genes (Dorca-Fornell et al., 2011). *SOC1* plays a central role in the flowering pathway (Moon et al., 2003; Yoo et al., 2005), and a recent study showed that *SOC1* also binds to its own locus and several flowering genes including the MADS-box members (Immink et al., 2012). Cis-motif 6 annotated as *SOC1* binding sites was found in the promoter of 15 *MIKC^C* genes (**Supplementary Table 7**), perhaps *SOC1* class members regulate other MADS-box genes during the longan flowering process.

Functional Conservation of ABCDE Model Genes in Longan

Longan is monecious and very little is known on longan floral biology. We collected four whorls of organs and ovaries from male and female flowers to predict the ABCDE model of longan (**Figure 6**). The expression of A-class *DlAP1* was notably higher in sepal than in other floral organs indicating the sepal identity of A function. Its close paralogs *DlFUL* was mainly expressed in the ovary. In Arabidopsis, *FUL* is required for valve differentiation and expansion in the ovary after fertilization (Zhang Y. et al., 2016). Generally, B class genes were restricted to expressed in the second and third whorls in model plants (Jack et al., 1992; Guo et al., 2015). In our results, *DlPI* was only detected in petals and stamens, which is in line with its expected function in petal and stamen identity specification. *DlAP3-1* was also expressed in the other two organs (sepal and carpel), indicating neo-functionalization in longan after the divergence of *PI* and *AP3*. Regarding the C class, *AG* was the first floral organs identity gene found to determine the stamen and carpel in Arabidopsis (Yanofsky et al., 1990). *DlAG* displayed a higher expression level in the stamen and carpel than the other organs, suggesting that *DlAG* maintains the C function of specifying both male and female reproductive organs in longan. A recent study revealed that Arabidopsis *AG* regulated sepal senescence (Jibran et al., 2017). In this study, the transcripts of *DlAG* were also found in sepals, supporting a potential role for sepal senescence in longan. Both *DlSHP* and *DlSTK* were relatively highly expressed in the ovary, we, therefore, hypothesized that *DlSHP* and *DlSTK* perform class D functions in longan ovule development.

All four members (*SEP1/2/3/4*) play redundant functions in determining floral organ identity in Arabidopsis (Zahn et al., 2006). A similar result was found in longan with four *DlSEP* having different expression profiles in tested organs. *DlSEP1* and 3 seem to account for a major position of E function due to their wider and higher expression level. *DlSEP4* might regulate the ovule development together with *DlFUL*, *DlSHP,* and *DlSTK* in the early stage. *AGL6* class was closely related to *SEP* in the phylogeny (**Figure 2**; Becker and Theissen, 2003), and orthologs of *AGL6* in rice, maize, and petunia have been regarded as regulators for multiple floral organ identity like Class E function (Ohmori et al., 2009; Rijpkema et al., 2009). Our analysis showed that *DlAGL6* possibly controls sepal, petal, and stamen of longan, not like *TaAGL6* which is required for all four whorls of wheat floral organs (Kong et al., 2021). Overall, the expression patterns of ABCDE model genes in longan are similar to typical models in Eudicots (**Figure 6D**).

The MADS Genes Are Involved in Flowering and Respond to Potassium Chlorate

MADS-box genes play significant roles in flowering processes in plants. Expression analysis allowed us to predict the function of longan MADS-box genes. Our results revealed a diverse expression pattern in different longan tissues. Compared to previous studies (Kofuji et al., 2003), the expression of most genes in type I and *MIKC** was not detected, with their function remaining unclear and it is possible they are only be expressed in specific cells or under specific conditions. Nevertheless, we found several genes such as *DlMADS21* and *DlMADS53* in root and pulp, suggesting these subfamilies with few studies never play a negligible role (**Supplementary Figure 5**). These type I and *MIKC** genes should be investigated in future work.

Flowering is the key factor in yield determination (Bangerth, 2009), but the problem of poor production persisted for a long time owing to low flowering rates in longan. Normally, flower induction in longan needs a period of low temperature (Suttitanawat et al., 2012), and longan can respond to $KClO_3$ for floral bud formation. The molecular mechanism associated with off-season flowering induction is of great value in scientific research and is of practical significance.

MIKC genes are of vital importance in reproductive development, especially the well know ABCDE model (Coen and Meyerowitz, 1991). In the present study, many *MIKC* genes seem to involve in the longan flowering process, as highly expressed genes usually play important roles in plant development. For example, *DlAP1*, *DlPI*, *DlAGL6* were highly expressed specifically in flower buds and flowers (**Figure 7**), which highlights their key roles in flowering. *DlAG, DlSHP, DlFUL, DlSEP1/2/3* showed high expression levels in multiple

tissues, inferring their function in regions beyond just flowers. More importantly, 11 *DlMADS* were up-regulated after $KClO_3$ treatment (**Figure 8A**), highlighting the involvement of MADS-box genes in off-season FI of longan.

Normally, *LFY* was considered as a determinant of flower initiation that plays a central role in flower pathways (Fornara et al., 2010). Recently, *LFY* was confirmed as a pioneer transcription factor to promote floral transition via upregulating *AP1* (Jin et al., 2021). However, the ortholog of *LFY* in longan showed no apparent changes during flower induction in "SJ" (Jue et al., 2019). On further investigation, we found the expression of *DlAP1* in "SJ" was significantly lower than in "SX" especially in the floral primordia stage, even though it was induced during floral transition in both accessions (**Supplementary Figure 8**). This result implied that the molecular pathways of flowering in "SJ" may differ from those in the other longan accessions. One A-class gene, *DlMADS98*, showed a higher expression level in "SJ." It is likely that the *DlMADS98* displayed function complementation with *DlAP1* in off-season FI.

In Arabidopsis, *FLC* and its five homologs, *MAF1-5* co-regulates flowering time by temperature-dependent alternative splicing (Scortecci et al., 2001; Bastow et al., 2004). Flowering induction in longan also requires low temperature (vernalization), thus, *FLC* might be functionally constrained in longan. *DlFLC* showed extensive-expression especially in longan vegetative and reproductive tissues and slightly decreased during off-season flowering. Significantly, the expression of *DlFLC* was lower in "SJ" than that in "SX" during the whole stages (**Supplementary Figure 8**). Given that *FLC* is regarded as a repressor in different flowering pathways (Rouse et al., 2002), we could speculate that *DlFLC* may be an inhibitor of off-season FI in longan. *SVP* is another important gene in response to environmental temperature to regulate plant flowering. The *SVP* class showed obvious expansion and the collinearity analysis implied it experienced tandem gene duplications like the *DAM* genes in Rosaceae (**Supplementary Figure 3**). RNA-seq and qRT-PCR analysis showed that *SVP* class genes such as *DlMADS82* tend to be expressed in leaf tissue and dormant buds, and this gene was also suppressed by $KClO_3$ (**Figures 7**, **8**). This leads us to speculate that *SVP* class genes could regulate vegetative development and be responsible for apical bud formation in response to dormancy.

Moreover, *SOC1* responds to multiple floral induction pathways including photoperiod, vernalization, and autonomous (Lee et al., 2000). It acts downstream of *FLOWERING LOCUS T (FT)* and regulates the expression of *LFY* which promotes transcription of *AP1* (Immink et al., 2012). On the other hand, it was also identified as a *FLC* suppressor (Lee et al., 2000). Contrary to the previous study, *DlSOC1* exhibited little expression in flowers or floral buds and was down-regulated after $KClO_3$ treatment. It is also notable that *DlMADS111* (*SOC1* clade) and *DlSOC1* showed lower expression levels in "SJ." Whether *SOC1* clade genes have altered functions in longan remains to be elucidated. From these results, it seems that genes in *DlFLC*, *DlSVP*, and *DlSOC1* have similar functions in maintaining vegetative growth and inhibiting floral transition, with further studies being required to verify the function of these genes.

Potential Roles of Longan MADS Genes in Fruit Development

Longan fruit, which contains many nutrients such as vitamins, carbohydrates, mineral elements, and amino acids, plays an important role in human anti-cancer, anti-aging, and brain development (Rangkadilok et al., 2005). The growth and development of longan fruit usually take 100–150 days. During this period, there are obvious changes in fruit size, nutritional composition, sweetness, and flavor, which are regulated by external environmental conditions, plant hormones, and many genes. In the present study, 9 of the MADS-box genes were highly expressed in fruit (**Figure 9**). *FRUITFULL FUL* is known for its role in controlling flowering time, carpel identity, fruit development, and leaf morphology (Coen and Meyerowitz, 1991; Gu et al., 1998; Torti et al., 2012). Functional studies in multiple species indicated that *FUL* orthologs have a conserved function for regulating fruit development even in fruits with diverse morphologies and structures (Jaakola et al., 2010; Bemer et al., 2012; Pabón-Mora et al., 2012; Fujisawa et al., 2014). Not surprisingly, in longan, *DlFUL* showed extremely high expression levels in young fruit, pericarp, and flowers, suggesting that this gene also has a conserved function for regulating the flowering time and fruit development (**Figures 7**, **9**).

Recent research showed that ectopic expression of *MdPI* controls apple fruit tissue growth and shape, indicating that the functions of B class genes are not limited to the development of petals and stamen (Yao et al., 2018). Analysis of *Sl-AGL11* overexpression in tomatoes revealed that the D class gene could regulate early fleshy fruit development (Huang et al., 2017). In this study, two *DlAP3* genes were not only expressed in flowers, but also expressed in fruit (seed, pericarp, and pulp) with different patterns during fruit development. The transcript of *DlSTK* has detected all the examined fruit organs with high levels in the pulp. It is likely that longan B and D class genes experienced different evolutionary fates from Arabidopsis and played functions for fruit development. In addition, a preview study has reported that *LeMADS-RIN* is necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus (Vrebalov et al., 2002). Arabidopsis *SEP* genes were the homologs of *LeMADS-RIN*. We determined that *DlSEP1* and *2* were extremely highly expressed with expression tending to increase during fruit development, which suggests *DlSEP1* and *2* may act as regulators of longan fruit ripening, with further studies being required to verify their function. It is noteworthy that a type I gene *DlMADS53* only expressed in the pulp of longan and remained at high expression levels during fruit ripening. To the best of our knowledge, there is no study about the orthologs of *DlMADS53* in plants. This result provides a new clue for the functional study of the MADS-box in the plants.

CONCLUSION

This study identified 114 MADS-box genes including 63 types I and 51 types II genes in the longan genome. Thirteen subfamilies

of *MIKC* genes were identified through phylogenetic analysis and the main difference between $MIKC^C$ and $MIKC^*$ proteins were the K-box domain. Analysis of cis-elements revealed that *Dof* transcription factors might directly regulate the $MIKC^C$ genes. Phylogeny and gene expressional analysis showed that the composition and expression of the ABCDE genes were conserved in longan. We also provided expression information for *DlMADS* in vegetative and reproductive tissues and this data led us to conclude that $MIKC^C$ genes play crucial roles in flowering. Several genes such as *DlSTK*, *DlSEP1/2,* and *DlMADS53* could be involved in fruit growth and ripening. In summary, this comprehensive analysis provided the basic resources to examine the molecular regulation of MADS-box genes in the longan reproduction process.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA741049.

AUTHOR CONTRIBUTIONS

JZ conceived the study and designed the experiments. BW, WH, YF, XF, JF, TZ, SZ, RM, and JZ carried out the experiments and analyzed the data. BW and JZ wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was financially supported by the National Key Research and Developmental Program of China (2018YFD1000104), the Project of Fujian Academy of Agricultural Sciences (AB2017-4), and the National Natural Science Foundation of China (31701884).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.813798/full#supplementary-material

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

Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., et al. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5328–5333. doi: 10.1073/pnas.97.10.5328

Aono, N., and Hasebe, M. (2006). Functional analysis of MADS-box genes in *Physcomitrella patens*. *Plant Cell Physiol. Suppl.* 2006, 498–498. doi: 10.14841/jspp.2006.0.498.0

Arora, R., Agarwal, P., Ray, S., Singh, A. K., Singh, V. P., Tyagi, A. K., et al. (2007). MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genom.* 8:242. doi: 10.1186/1471-2164-8-242

Bangerth, K. F. (2009). Floral induction in mature, perennial angiosperm fruit trees: similarities and discrepancies with annual/biennial plants and the involvement of plant hormones. *Sci. Hortic.* 122, 153–163. doi: 10.1016/j.scienta.2009.06.014

Bastow, R., Mylne, J. S., Lister, C., Lippman, Z., Martienssen, R. A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427, 164–167. doi: 10.1038/nature02269

Becker, A., and Theissen, G. (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Mol. Phylogenet. Evol.* 29, 464–489. doi: 10.1016/s1055-7903(03)00207-0

Bemer, M., Karlova, R., Ballester, A. R., Tikunov, Y. M., Bovy, A. G., Wolters-Arts, M., et al. (2012). The tomato FRUITFULL homologs TDR4/FUL1 and MBP7/FUL2 regulate ethylene-independent aspects of fruit ripening. *Plant Cell* 24, 4437–4451. doi: 10.1105/tpc.112.103283

Bielenberg, D. G., Wang, Y., Li, Z., Zhebentyayeva, T., Fan, S., Reighard, G. L., et al. (2008). Sequencing and annotation of the evergrowing locus in peach [*Prunus persica* (L.) Batsch] reveals a cluster of six MADS-box transcription factors as candidate genes for regulation of terminal bud formation. *Tree Genet. Genom.* 4, 495–507. doi: 10.1007/s11295-007-0126-9

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Chang, Q. (2010). *Study on Relationship Between Off-Season Floral Induction in Longan and Nutrition of the Carbon and Nitrogen*. Master's thesis. Fuzhou: Fujian Agriculture & Forestry University.

Chen, S., Liu, H., Chen, W., Xie, D., and Zheng, S. (2009). Proteomic analysis of differentially expressed proteins in longan flowering reversion buds. *Sci. Hortic.* 122, 275–280. doi: 10.1016/j.scienta.2009.05.015

Chen, W., Chao, G., and Singh, K. B. (1996). The promoter of a H2O2-inducible, *Arabidopsis* glutathione S-transferase gene contains closely linked OBF- and OBP1-binding sites. *Plant J.* 10, 955–966. doi: 10.1046/j.1365-313x.1996.10060955.x

Chen, X., Wang, D., Liu, C., Wang, M., Wang, T., Zhao, Q., et al. (2012). Maize transcription factor Zmdof1 involves in the regulation of Zm401 gene. *Plant Growth Regul.* 66, 271–284. doi: 10.1007/s10725-011-9651-5

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

De Bodt, S., Raes, J., Van de Peer, Y., and Theissen, G. (2003). And then there were many: MADS goes genomic. *Trends Plant Sci.* 8, 475–483. doi: 10.1016/j.tplants.2003.09.006

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M. F. (2004). The SEP4 gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. *Curr. Biol.* 14, 1935–1940. doi: 10.1016/j.cub.2004.10.028

Dorca-Fornell, C., Gregis, V., Grandi, V., Coupland, G., Colombo, L., and Kater, M. M. (2011). The *Arabidopsis* SOC1-like genes AGL42, AGL71 and AGL72 promote flowering in the shoot apical and axillary meristems. *Plant J.* 67, 1006–1017. doi: 10.1111/j.1365-313X.2011.04653.x

Duan, W., Song, X., Liu, T., Huang, Z., Ren, J., Hou, X., et al. (2015). Genome-wide analysis of the MADS-box gene family in *Brassica rapa* (Chinese cabbage). *Mol. Genet. Genom.* 290, 239–255. doi: 10.1007/s00438-014-0912-7

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

Fan, C. M., Wang, X., Wang, Y. W., Hu, R. B., Zhang, X. M., Chen, J. X., et al. (2013). Genome-wide expression analysis of soybean MADS genes showing potential function in the seed development. *PLoS One* 8:e62288. doi: 10.1371/journal.pone.0062288

Fornara, F., de Montaigu, A., and Coupland, G. (2010). SnapShot: control of flowering in *Arabidopsis*. *Cell* 141, 550–550.e552. doi: 10.1016/j.cell.2010.04.024

Fujisawa, M., Shima, Y., Nakagawa, H., Kitagawa, M., Kimbara, J., Nakano, T., et al. (2014). Transcriptional regulation of fruit ripening by tomato FRUITFULL

homologs and associated MADS box proteins. *Plant Cell* 26, 89–101. doi: 10.1105/tpc.113.119453

Goto, K., and Meyerowitz, E. M. (1994). Function and regulation of the *Arabidopsis* floral homeotic gene PISTILLATA. *Genes Dev.* 8, 1548–1560. doi: 10.1101/gad.8.13.1548

Gross, T., Broholm, S., and Becker, A. (2018). CRABS CLAW acts as a bifunctional transcription factor in flower development. *Front. Plant Sci.* 9:835. doi: 10.3389/fpls.2018.00835

Gu, Q., Ferrándiz, C., Yanofsky, M. F., and Martienssen, R. (1998). The FRUITFULL MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. *Development* 125, 1509–1517.

Guo, S., Sun, B., Looi, L. S., Xu, Y., Gan, E. S., Huang, J., et al. (2015). Co-ordination of flower development through epigenetic regulation in two model species: rice and *Arabidopsis*. *Plant Cell Physiol.* 56, 830–842. doi: 10.1093/pcp/pcv037

Harrington, M., Edwards, K., Johnson, S., Chase, M., and Gadek, P. (2005). Phylogenetic inference in *Sapindaceae sensu* lato using plastid matK and rbcL DNA sequences. *Syst. Bot.* 30, 366–382. doi: 10.1600/0363644054223549

Hepworth, S. R., Klenz, J. E., and Haughn, G. W. (2006). UFO in the *Arabidopsis* inflorescence apex is required for floral-meristem identity and bract suppression. *Planta* 223, 769–778. doi: 10.1007/s00425-005-0138-3

Hou, X.-J., Liu, S.-R., Khan, M. R. G., Hu, C.-G., and Zhang, J.-Z. (2013). Genome-wide identification, classification, expression profiling, and SSR marker development of the MADS-Box gene family in citrus. *Plant Mol. Biol. Report.* 32, 28–41. doi: 10.1007/s11105-013-0597-9

Huang, B., Routaboul, J. M., Liu, M., Deng, W., Maza, E., Mila, I., et al. (2017). Overexpression of the class D MADS-box gene Sl-AGL11 impacts fleshy tissue differentiation and structure in tomato fruits. *J. Exp. Bot.* 68, 4869–4884. doi: 10.1093/jxb/erx303

Imaizumi, T., Schultz, T. F., Harmon, F. G., Ho, L. A., and Kay, S. A. (2005). FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in *Arabidopsis*. *Science* 309, 293–297. doi: 10.1126/science.1110586

Immink, R. G. H., Posé, D., Ferrario, S., Ott, F., Kaufmann, K., Valentim, F. L., et al. (2012). Characterization of SOC1's central role in flowering by the identification of its upstream and downstream regulators. *Plant Physiol.* 160, 433–449. doi: 10.1104/pp.112.202614

Jaakola, L., Poole, M., Jones, M. O., Kämäräinen-Karppinen, T., Koskimäki, J. J., Hohtola, A., et al. (2010). A SQUAMOSA MADS box gene involved in the regulation of anthocyanin accumulation in bilberry fruits. *Plant Physiol.* 153, 1619–1629. doi: 10.1104/pp.110.158279

Jack, T., Brockman, L. L., and Meyerowitz, E. M. (1992). The homeotic gene APETALA3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* 68, 683–697. doi: 10.1016/0092-8674(92)90144-2

Jibran, R., Tahir, J., Cooney, J., Hunter, D. A., and Dijkwel, P. P. (2017). *Arabidopsis* AGAMOUS regulates sepal senescence by driving jasmonate production. *Front. Plant Sci.* 8:2101. doi: 10.3389/fpls.2017.02101

Jin, R., Klasfeld, S., Zhu, Y., Fernandez Garcia, M., Xiao, J., Han, S.-K., et al. (2021). LEAFY is a pioneer transcription factor and licenses cell reprogramming to floral fate. *Nat. Commun.* 12:626. doi: 10.1038/s41467-020-20883-w

Jofuku, K. D., den Boer, B. G., Van Montagu, M., and Okamuro, J. K. (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* 6, 1211–1225. doi: 10.1105/tpc.6.9.1211

Johansen, B., Pedersen, L. B., Skipper, M., and Frederiksen, S. (2002). MADS-box gene evolution-structure and transcription patterns. *Mol. Phylogenet. Evol.* 23, 458–480. doi: 10.1016/s1055-7903(02)00032-5

Jue, D., Sang, X., Liu, L., Shu, B., Wang, Y., Liu, C., et al. (2019). Comprehensive analysis of the longan transcriptome reveals distinct regulatory programs during the floral transition. *BMC Genom.* 20:126. doi: 10.1186/s12864-019-5461-3

Kania, T., Russenberger, D., Peng, S., Apel, K., and Melzer, S. (1997). FPF1 promotes flowering in *Arabidopsis*. *Plant Cell* 9, 1327–1338. doi: 10.1105/tpc.9.8.1327

Kaufmann, K., Melzer, R., and Theissen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene* 347, 183–198. doi: 10.1016/j.gene.2004.12.014

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kofuji, R., Sumikawa, N., Yamasaki, M., Kondo, K., Ueda, K., Ito, M., et al. (2003). Evolution and divergence of the MADS-box gene family based on genome-wide expression analyses. *Mol. Biol. Evol.* 20, 1963–1977. doi: 10.1093/molbev/msg216

Kong, X., Wang, F., Geng, S., Guan, J., Tao, S., Jia, M., et al. (2021). The wheat AGL6-like MADS-box gene is a master regulator for floral organ identity and a target for spikelet meristem development manipulation. *Plant Biotechnol. J.* 20, 75–88. doi: 10.1111/pbi.13696

Koo, S. C., Bracko, O., Park, M. S., Schwab, R., Chun, H. J., Park, K. M., et al. (2010). Control of lateral organ development and flowering time by the *Arabidopsis thaliana* MADS-box Gene AGAMOUS-LIKE6. *Plant J.* 62, 807–816. doi: 10.1111/j.1365-313X.2010.04192.x

Kooiker, M., Airoldi, C. A., Losa, A., Manzotti, P. S., Finzi, L., Kater, M. M., et al. (2005). BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene SEEDSTICK. *Plant Cell* 17, 722–729. doi: 10.1105/tpc.104.030130

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., and Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* 20:278. doi: 10.1186/s13059-019-1910-1

Krolikowski, K. A., Victor, J. L., Wagler, T. N., Lolle, S. J., and Pruitt, R. E. (2003). Isolation and characterization of the *Arabidopsis* organ fusion gene HOTHEAD. *Plant J.* 35, 501–511. doi: 10.1046/j.1365-313X.2003.01824.x

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lai, Z.-X., Chen, C. L., Zeng, L. H., and Chen, Z. G. (2000). Somatic embryogenesis in longan (*Dimocarpus longan* Lour.). *Somat. Embryogen. Woody Plants* 6, 415–431. doi: 10.1007/978-94-017-3030-3_13

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Lee, H., Suh, S. S., Park, E., Cho, E., Ahn, J. H., Kim, S. G., et al. (2000). The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. *Genes Dev.* 14, 2366–2376. doi: 10.1101/gad.813600

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* 61, 2247–2254. doi: 10.1093/jxb/erq098

Li, D., Yang, C., Li, X., Gan, Q., Zhao, X., and Zhu, L. (2009). Functional characterization of rice OsDof12. *Planta* 229, 1159–1169. doi: 10.1007/s00425-009-0893-7

Lin, Y., Min, J., Lai, R., Wu, Z., Chen, Y., Yu, L., et al. (2017). Genome-wide sequencing of longan (*Dimocarpus longan* Lour.) provides insights into molecular basis of its polyphenol-rich characteristics. *Gigascience* 6, 1–14. doi: 10.1093/gigascience/gix023

Liu, J., Zhang, J., Zhang, J., Miao, H., Wang, J., Gao, P., et al. (2017). Genome-wide analysis of banana MADS-box family closely related to fruit development and ripening. *Sci. Rep.* 7:3467. doi: 10.1038/s41598-017-03897-1

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., et al. (2005). Modeling gene and genome duplications in eukaryotes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5454–5459. doi: 10.1073/pnas.0501102102

Manochai, P., Sruamsiri, P., Wiriya-alongkorn, W., Naphrom, D., Hegele, M., and Bangerth, F. (2005). Year around off season flower induction in longan (*Dimocarpus longan*, Lour.) trees by KClO3 applications: potentials and problems. *Sci. Hortic.* 104, 379–390. doi: 10.1016/j.scienta.2005.01.004

Masiero, S., Colombo, L., Grini, P. E., Schnittger, A., and Kater, M. M. (2011). The emerging importance of type I MADS box transcription factors for plant reproduction. *Plant Cell* 23, 865–872. doi: 10.1105/tpc.110.081737

Mei, Z. Q., Fu, S. Y., Yu, H. Q., Yang, L. Q., Duan, C. G., Liu, X. Y., et al. (2014). Genetic characterization and authentication of *Dimocarpus longan* Lour. using an improved RAPD technique. *Genet. Mol. Res.* 13, 1447–1455. doi: 10.4238/2014.March.6.3

Meister, R. J., Williams, L. A., Monfared, M. M., Gallagher, T. L., Kraft, E. A., Nelson, C. G., et al. (2004). Definition and interactions of a positive regulatory element of the *Arabidopsis* INNER NO OUTER promoter. *Plant J.* 37, 426–438. doi: 10.1046/j.1365-313X.2003.01971.x

Messenguy, F., and Dubois, E. (2003). Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene* 316, 1–21. doi: 10.1016/s0378-1119(03)00747-9

Moon, J., Suh, S. S., Lee, H., Choi, K. R., Hong, C. B., Paek, N. C., et al. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*. *Plant J.* 35, 613–623. doi: 10.1046/j.1365-313x.2003.01833.x

Moore, S., Vrebalov, J., Payton, P., and Giovannoni, J. (2002). Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato. *J. Exper. Bot.* 53, 2023–2030. doi: 10.1093/jxb/erf057

Ohmori, S., Kimizu, M., Sugita, M., Miyao, A., Hirochika, H., Uchida, E., et al. (2009). MOSAIC FLORAL ORGANS1, an AGL6-like mads box gene, regulates floral organ identity and meristem fate in rice. *Plant Cell* 21, 3008–3025. doi: 10.1105/tpc.109.068742

Pabón-Mora, N., Ambrose, B. A., and Litt, A. (2012). Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth identity, and fruit development. *Plant Physiol.* 158, 1685–1704. doi: 10.1104/pp.111.192104

Parenicova, L., de Folter, S., Kieffer, M., Horner, D. S., Favalli, C., Busscher, J., et al. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. *Plant Cell* 15, 1538–1551. doi: 10.1105/tpc.011544

Park, D. H., Lim, P. O., Kim, J. S., Cho, D. S., Hong, S. H., and Nam, H. G. (2003). The *Arabidopsis* COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. *Plant J.* 34, 161–171. doi: 10.1046/j.1365-313x.2003.01710.x

Passmore, S., Maine, G. T., Elble, R., Christ, C., and Tye, B. K. (1988). Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary for mating of MAT alpha cells. *J. Mol. Biol.* 204, 593–606. doi: 10.1016/0022-2836(88)90358-0

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405, 200–203. doi: 10.1038/35012103

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

Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077

Project, A. G. (2013). The Amborella genome and the evolution of flowering plants. *Science* 342:1241089. doi: 10.1126/science.1241089

Rangkadilok, N., Worasuttayangkurn, L., Bennett, R. N., and Satayavivad, J. (2005). Identification and quantification of polyphenolic compounds in longan (*Euphoria longana* Lam.) fruit. *J. Agric. Food Chem.* 53, 1387–1392. doi: 10.1021/jf0403484

Riechmann, J. L., and Meyerowitz, E. M. (1997). MADS domain proteins in plant development. *Biol. Chem.* 378, 1079–1101.

Rijpkema, A. S., Zethof, J., Gerats, T., and Vandenbussche, M. (2009). The petunia AGL6 gene has a SEPALLATA-like function in floral patterning. *Plant J.* 60, 1–9. doi: 10.1111/j.1365-313X.2009.03917.x

Rouse, D. T., Sheldon, C. C., Bagnall, D. J., Peacock, W. J., and Dennis, E. S. (2002). FLC, a repressor of flowering, is regulated by genes in different inductive pathways. *Plant J.* 29, 183–191. doi: 10.1046/j.0960-7412.2001.01210.x

Scortecci, K. C., Michaels, S. D., and Amasino, R. M. (2001). Identification of a MADS-box gene, FLOWERING LOCUS M, that represses flowering. *Plant J.* 26, 229–236. doi: 10.1046/j.1365-313x.2001.01024.x

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Simonini, S., Roig-Villanova, I., Gregis, V., Colombo, B., Colombo, L., and Kater, M. M. (2012). Basic pentacysteine proteins mediate MADS domain complex binding to the DNA for tissue-specific expression of target genes in *Arabidopsis*. *Plant Cell* 24, 4163–4172. doi: 10.1105/tpc.112.103952

Solovyev, V., Kosarev, P., Seledsov, I., and Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* 7:S10. doi: 10.1186/gb-2006-7-s1-s10

Sommer, H., Beltrán, J. P., Huijser, P., Pape, H., Lönnig, W. E., Saedler, H., et al. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. *EMBO J.* 9, 605–613.

Sringarm, K., Potchanasin, P., Sruamsiri, P., and Bangerth, K. F. (2009). Floral induction (FI) in longan (*Dimocarpus longan*, Lour.) trees—The possible participation of endogenous hormones: II. Low temperature and potassium chlorate effects on hormone concentrations in and their export out of leaves. *Sci. Hortic.* 122, 295–300. doi: 10.1016/j.scienta.2008.11.031

Subhadrabandhu, S., and Yapwattanaphun, C. (2001). Regulation off-season flowering of longan in Thailand. *Acta Hortic.* 558, 193–198. doi: 10.17660/ActaHortic.2001.558.26

Suttitanawat, P., Sruamsiri, P., and Sringarm, K. (2012). Changes in cytokinins concentrations during induction period of longan cv. Daw in sand culture. *J. Agric. Technol.* 8, 2353–2362.

Theissen, G. (2001). Development of floral organ identity: stories from the MADS house. *Curr. Opin. Plant Biol.* 4, 75–85. doi: 10.1016/s1369-5266(00)00139-4

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J. T., Münster, T., et al. (2000). A short history of MADS-box genes in plants. *Plant Mol. Biol.* 42, 115–149.

Theissen, G., Kim, J. T., and Saedler, H. (1996). Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. *J. Mol. Evol.* 43, 484–516. doi: 10.1007/bf02337521

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

Torti, S., Fornara, F., Vincent, C., Andrés, F., Nordström, K., Göbel, U., et al. (2012). Analysis of the *Arabidopsis* shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a Leucine-rich repeat protein that promotes flowering. *Plant Cell* 24, 444–462. doi: 10.1105/tpc.111.092791

Van de Peer, Y. (2004). Computational approaches to unveiling ancient genome duplications. *Nat. Rev. Genet.* 5, 752–763. doi: 10.1038/nrg1449

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., et al. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* 296, 343–346. doi: 10.1126/science.1068181

Wang, L., Yin, X., Cheng, C., Wang, H., Guo, R., Xu, X., et al. (2015). Evolutionary and expression analysis of a MADS-box gene superfamily involved in ovule development of seeded and seedless grapevines. *Mol. Genet. Genom.* 290, 825–846. doi: 10.1007/s00438-014-0961-y

Wang, Y., Tang, H., DeBarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293

Ward, J. M., Cufr, C. A., Denzel, M. A., and Neff, M. M. (2005). The Dof transcription factor OBP3 modulates phytochrome and cryptochrome signaling in *Arabidopsis*. *Plant Cell* 17, 475–485. doi: 10.1105/tpc.104.027722

Winterhagen, P., Tiyayon, P., Samach, A., Hegele, M., and Wünsche, J. N. (2013). Isolation and characterization of FLOWERING LOCUS T subforms and APETALA1 of the subtropical fruit tree *Dimocarpus longan*. *Plant Physiol. Biochem.* 71, 184–190. doi: 10.1016/j.plaphy.2013.07.013

Yanagisawa, S. (2002). The Dof family of plant transcription factors. *Trends Plant Sci.* 7, 555–560. doi: 10.1016/S1360-1385(02)02362-2

Yanagisawa, S., and Schmidt, R. J. (1999). Diversity and similarity among recognition sequences of Dof transcription factors. *Plant J.* 17, 209–214. doi: 10.1046/j.1365-313X.1999.00363.x

Yang, Y., Fanning, L., and Jack, T. (2003). The K domain mediates heterodimerization of the *Arabidopsis* floral organ identity proteins, APETALA3 and PISTILLATA. *Plant J.* 33, 47–59. doi: 10.1046/j.0960-7412.2003.01473.x

Yanofsky, M. F., Ma, H., Bowman, J. L., Drews, G. N., Feldmann, K. A., and Meyerowitz, E. M. (1990). The protein encoded by the *Arabidopsis* homeotic gene agamous resembles transcription factors. *Nature* 346, 35–39. doi: 10.1038/346035a0

Yao, J.-L., Xu, J., Tomes, S., Cui, W., Luo, Z., Deng, C., et al. (2018). Ectopic expression of the PISTILLATA homologous MdPI inhibits fruit tissue growth and changes fruit shape in apple. *Plant J.* 2:e00051. doi: 10.1002/pld3.51

Yoo, S. K., Chung, K. S., Kim, J., Lee, J. H., Hong, S. M., Yoo, S. J., et al. (2005). CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in *Arabidopsis*. *Plant Physiol.* 139, 770–778. doi: 10.1104/pp.105.066928

Zahn, L. M., Feng, B., and Ma, H. (2006). Beyond the ABC-model: regulation of floral homeotic genes. *Adv. Bot. Res.* 44, 163–207.

Zhang, H. N., Shi, S. Y., Li, W. C., Shu, B., Liu, L. Q., Xie, J. H., et al. (2016). Transcriptome analysis of 'Sijihua' longan (*Dimocarpus longan* L.) based on next-generation sequencing technology. *J. Hortic. Sci. Biotechnol.* 91, 180–188. doi: 10.1080/14620316.2015.1133539

Zhang, Y., Shen, Y. Y., Wu, X. M., and Wang, J. B. (2016). The basis of pod dehiscence: anatomical traits of the dehiscence zone and expression of eight pod shatter-related genes in four species of *Brassicaceae*. *Biol. Plant.* 60, 343–354. doi: 10.1007/s10535-016-0599-1

Zhang, L., Chen, F., Zhang, X., Li, Z., Zhao, Y., Lohaus, R., et al. (2020). The water lily genome and the early evolution of flowering plants. *Nature* 577, 79–84. doi: 10.1038/s41586-019-1852-5

Zhao, Y., Broholm, S. K., Wang, F., Rijpkema, A. S., Lan, T., Albert, V. A., et al. (2020). TCP and MADS-Box transcription factor networks regulate heteromorphic flower type identity in *Gerbera hybrida*. *Plant Physiol.* 184, 1455–1468. doi: 10.1104/pp.20.00702

Zhao, Y., Li, X., Chen, W., Peng, X., Cheng, X., Zhu, S., et al. (2011). Whole-genome survey and characterization of MADS-box gene family in maize and sorghum. *Plant Cell Tissue Organ Cult.* 105, 159–173. doi: 10.1007/s11240-010-9848-8

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fpls.2021.821147

Investigating the Mechanism of Unilateral Cross Incompatibility in Longan (*Dimocarpus longan* Lour.) Cultivars (Yiduo × Shixia)

Jing Wang[1,2], Ji Chen[3], Shilian Huang[1,2], Dongmei Han[1,2], Jianguang Li[1,2] and Dongliang Guo[1,2]**

[1] Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences, Guangzhou, China, [2] Key Laboratory of South Subtropical Fruit Biology and Genetic Resource Utilization, Guangdong Provincial Key Laboratory of Tropical and Subtropical Fruit Tree Research, Ministry of Agriculture and Rural Affairs, Guangzhou, China, [3] College of Horticulture, South China Agricultural University, Guangzhou, China

OPEN ACCESS

Edited by:
Yuexue Liu, Shenyang Agricultural University, China

Reviewed by:
Gaetano Distefano, University of Catania, Italy
Maria Jazmin Abraham-Juarez, National Laboratory of Genomics for Biodiversity, Center for Research and Advanced Studies, National Polytechnic Institute of Mexico (CINVESTAV), Mexico

***Correspondence:**
Jianguang Li
lijianguang@gdaas.cn
Dongliang Guo
guodongliang@gdaas.cn

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

Received: *23 November 2021*
Accepted: *30 December 2021*
Published: *11 February 2022*

Citation:
Wang J, Chen J, Huang S, Han D, Li J and Guo D (2022) Investigating the Mechanism of Unilateral Cross Incompatibility in Longan (Dimocarpus longan Lour.) Cultivars (Yiduo × Shixia). Front. Plant Sci. 12:821147. doi: 10.3389/fpls.2021.821147

Longan (*Dimocarpus longan* Lour.) is an important subtropical fruit tree in China. Nearly 90% of longan fruit imports from Thailand are from the cultivar Yiduo. However, we have observed that there exists a unilateral cross incompatibility (UCI) when Yiduo is used as a female parent and Shixia (a famous Chinese cultivar) as a male parent. Here, we performed a comparative transcriptome analysis coupled with microscopy of pistils from two reciprocal pollination combinations [Shixia♂ × Yiduo♀(SY) and Yiduo♀ × Shixia♂(YS)] 4, 8, 12, and 24 h after pollination. We also explored endogenous jasmonic acid (JA) and jasmonyl isoleucine (JA-Ile) levels in pistils of the crosses. The microscopic observations showed that the UCI was sporophytic. The endogenous JA and JA-Ile levels were higher in YS than in SY at the studied time points. We found 7,251 differentially expressed genes from the transcriptome analysis. Our results highlighted that genes associated with JA biosynthesis and signaling, pollen tube growth, cell wall modification, starch and sucrose biosynthesis, and protein processing in endoplasmic reticulum pathways were differentially regulated between SY and YS. We discussed transcriptomic changes in the above-mentioned pathways regarding the observed microscopic and/or endogenous hormone levels. This is the first report on the elaboration of transcriptomic changes in longan reciprocal pollination combination showing UCI. The results presented here will enable the longan breeding community to better understand the mechanisms of UCI.

Keywords: Chinese longan, intraspecific cross incompatibility, jasmonic acid metabolism, plant-hormone, gene expression

INTRODUCTION

Longan (*Dimocarpus longan* Lour.) belongs to the *Sapindaceae* family. It is one of the most important subtropical fruit trees indigenous in China. Within China, major longan-producing areas are distributed between 18–31.16°N latitude and 100°44′–122°E longitude, Guanxi, Guangdong, and Fujian are the major production areas (Qiu, 2012). The fruit is rich in nutrients and has been regarded as "precious tonic" since ancient times. It is believed that longan originated in China,

since it has been cultivated for thousands of years (Subhadrabandhu and Yapwattanaphun, 2001). Recently, China has been importing large volumes of longan from nearby countries, i.e., Thailand, Vietnam, and Cambodia (Hasachoo and Kalaya, 2013). In 2020, China imported 72.67 MMT of longan fruits[1]. The increasing imports are impacting local production. There are more than 400 longan cultivars registered in China. The major cultivars are Shixia, Chuliang, Dawuyuan, and Fuyan. These cultivars are cultivated in different regions depending on their traits and suitability to the local climate. Shixia is the most cultivated cultivar in Guangdong province. However, its production is lower than those that are grown in Thailand. A major example is the Thai longan cultivar "E-Daw," which is known as Yiduo in China (Choo, 2000). Yiduo is a tropical ecotype that was introduced in China by Thailand. Its yield is relatively higher than that the Shixia or other cultivars in China. Yiduo production per mu can reach 2,000–3,000 kg in Thailand. After its introduction in China, its production has declined and is mainly dependent on self-pollination to bear fruits. Yiduo has several useful traits such as easy flowering, coarse and large fruits, vigorous growth, high edible rate of fruit, and high yield and quality of fruits, which have attracted the attention of Chinese longan breeders to include it in Chinese longan breeding programs. Our research group has tried to improve Yiduo using local cultivars as male parent by hybridization. However, in three successive years, we have noticed that there exists a unilateral cross incompatibility (UCI) when Yiduo is used as a female parent and Shixia as a male parent, i.e., Yiduo♀ × Shixia ♂(YS). Our earlier data showed a fruit setting rate of 22.4% and zero percent in SY and YS, respectively. Understanding this reproductive barrier is a prime target for longan breeders to uplift the quality of longan in China.

Unilateral cross incompatibility has been reported in multiple plant species, e.g., capsicum (Onus and Pickersgill, 2004), faba bean (Abdalla, 1977), and field mustard (Takada et al., 2017), and is defined as an intraspecific relationship in which pollinations are only compatible in one direction (Tovar-Méndez et al., 2014). In *Solanum* species, UCI has been reported because of rejection of pollen from self-compatible species on the pistil of self-incompatible species (Tovar-Méndez et al., 2014). In tomato, this UCI was linked to the absence of S-RNAse. In brassica species, the self-incompatibility is regulated by an S-haplotype-specific interaction between S-receptor kinase (stigma-specific expression) and S locus protein 11 (tapetum-cell-specific expression) (Takada et al., 2017). Generally, it is considered that pre-fertilization/pre-zygotic/post-pollination barriers can be responsible for such UCI that could be due to failure of pollen germination or pollen growth in pistil (De Nettancourt, 2001). Pollen tube growth is different from that of other plant cells, as its growth is mainly restricted to the tip region (Mascarenhas, 1993). Under conditions of pollen competition, only fast-growing pollen tubes can accomplish effective fertilization. Pollen tube growth varies in different species (Stone et al., 2004), and certain factors such as calcium and potassium ion concentration (Caser, 2017), biosynthesis of cell wall polymers (Mollet et al., 2013), and availability and concentration of sucrose/carbohydrates (O'Kelley, 1955; Stadler et al., 1999) can contribute to pollen tube growth. These factors are controlled by a large number of genes belonging to different biosynthetic processes, i.e., plant hormone signaling, cell enlargement, metabolism-related pathways, cell wall biosynthesis and rearrangement, and protein processing (Becker and Feijó, 2007). An understanding of the expression of genes in specific pathways would enable us to explore the possible reasons for UCI in YS. Recent advancements in transcriptome sequencing technology have enabled researchers to understand the molecular mechanisms underlying cross incompatibility in maize (Wang et al., 2018), self-incompatibility in lemon (Zhang et al., 2015), tea (Zhang et al., 2016), and oilseed camelia (He et al., 2020), and pollen tube development in olive (Iaria et al., 2016). A similar approach would help us to explore transcriptomic signatures by comparing different pollination combinations in Yiduo and Shixia longan trees.

Methods to overcome UCI include genetic rescue (Holmes et al., 2008), use of hormone application together with irradiated pollen and mixed pollination, grafted ovary method, cut-style method, e.g., *Lycopersicon* (*Lycopersicum esculentum* × *Lycopersicum peruvianum*) (Majid, 1964), Tulipa (Van Creij et al., 1997), and lily (Van Creij et al., 2000). Since pollen germination and pollen tube growth are development-related processes, the role of particular hormones cannot be negated. For example, it has been reported that pollen germination and pollen tube growth in apricot and *Pinus nigra* are influenced by methyl jasmonate (MeJA) (Muradoğlu et al., 2010; Çetinbaş-Genç and Vardar, 2020). Similarly, higher endogenous jasmonic acid (JA) levels have also been related to limited pollen germination in Arabidopsis (Ju et al., 2016).

In this study, we designed a reciprocal pollination combination of these two cultivars and explored transcriptomic signatures in pistil in different hours after pollination (HAP) that could be associated with UCI. We discuss the possibilities

[1]www.tridge.com

Abbreviations: 12-OPDA, 12-13-epoxylinoleic acid to 12-oxophytodienoate; AA, α-amylase; AOC, allene oxide cyclase; AOS, allene oxide synthase; ARO1, armadillo repeat only 1; BA, β-amylase; BAK, BRI1-associated receptor kinase; BERUCTF, β-fructofuranosidase; BGL, β-glucosidase; BGLs, β-glucosidases; CA3OM, 3-O-methyltransferase; CAD, cinnamyl-alcohol dehydrogenase; CESA, cellulose synthase; CHI, chintinase; CHX15, cation/H(+) antiporter 15; CK, control; CNGC18, cyclic nucleotide-gated channel; COG, clusters of orthologous groups; COI, coronatine-insensitive protein; CPK, calcium dependent protein kinase; DAD, defective in anther dehiscence; DEGs, differentially expressed genes; EG, endoglucanase; ERF, ethylene responsive factor; FDR, false discovery rate; FPKM, fragments per kilobase of transcript per million fragments mapped; GAG, 1,4-alpha-galacturonidase; GH3, glycosylhydrolase family 3; GID, Gibberellin receptor GID; GIP1, glucose-1-phosphate adenylyltransferase; GO, gene ontology; HAP, hours after pollination; IAA, indole-3-acetic acid; INT13, myo-inositol transporter 13; JA, jasmonic acid; JA-Ile, jasmonyl isoleucine; JAR1, jasmonoyl-isoleucine synthetase; JAZ, jasmonate ZIM domain; JMT, jasmonic acid carboxyl methyltransferase; KEGG, Kyoto Encyclopedia of Genes and Genomes; LOX, lipoxigenase; MeJA, methyl jasmonate; MYC2, transcription factor MYC2; OPR, OPDA reductase; PAL, phenylalanine ammonia lyase; PCA, principal component Analysis; PCC, Pearson's correlation coefficient; PERK, proline-rich receptor-like protein kinases; PLA, phospholipase A1; PME, pectin methylesterase; PMEI, PME-inhibitor; PRP, proline rich protein; RALFL, RALF-like; S, Shixia; SAUR, small auxin up-regulated RNA; SPS, sucrose phosphate synthase; SRK, S-receptor kinase; SS13, strictosidine synthase-like 13; SuSy, sucrose synthase; SYP124, syntaxin-124; TGA, TGA transcription factor; UCI, unilateral cross incompatability; XETs, xyloglucan:xyloglucosyl transferases; XIPs, xylanase inhibitors; Y, Yiduo.

that differences in pollen tube growth can be linked with UCI. Furthermore, we tried to understand and explain if there are changes in the endogenous levels of JA and its derivative MeJA. We discussed the possible roles of pollen tube-related genes, phytohormone signaling, starch and sucrose biosynthesis, phenylpropanoid biosynthesis, JA metabolism and signaling, and protein processing in endoplasmic reticulum pathways in UCI.

MATERIALS AND METHODS

Plant Materials

Three plants of two *Dimocarpus longan* Lour. Cultivars, i.e., Shixia and Yiduo, were selected from the Longan Resource Nursery of Guangzhou City, Guangzhou, China. Care was taken while selecting the longan trees, and it was ensured that the selected trees were of the same growth and age (10 years old). The flower spikes present on different sides of the trees were chosen on 28 March 2020 for pollination and covered with waterproof sulfuric acid bags (40 cm^2 × 30 cm^2). Artificial pollinations were carried out in two reciprocal combinations, i.e., combination I: Yiduo♀ × Shixia♂(YS) and combination II: Shixia♀ × Yiduo♂(SY) (**Figure 1**). The pollinations were carried out as reported earlier (McConchie et al., 1994; Pham, 2012). Each pollination combination was marked for identification. Briefly, on 31 March 2020 at 8:30 a.m., the bags were removed at the same time from the flowers chosen as female, and pistils were removed before pollination and immediately stored in liquid nitrogen; each variety had three replicates. After this, artificial pollination was carried out using the male flowers bloomed in the same morning, and the pollinated flowers were bagged. Pistils were collected 0, 1, 4, 8, 12, and 24 h after pollination (HAP), quickly frozen in liquid nitrogen, and stored at −70° or in a Kano fixing solution. The stored samples were then used for detection of endogenous phytohormone levels (0, 4, 8, 12, and 24 HAP), transcriptome sequencing, and *in vitro* study on pollen germination and pollen tube growth (0, 4, 8, 12, and 24 HAP).

Determination of Phytohormone Levels

An AB Sciex QTRAP6500 LC-MS/MS system was used to detect the content of phytohormones before and after pollination in the different pollination combinations reported earlier (Xiao et al., 2018; López-Cristoffanini et al., 2019).

Study on Pollen Germination and Pollen Tube Growth

Pollen *in situ* germination and pollen tube growth in the style and ovary were observed as described by Xie et al. (2019), with slight modifications. Shixia pistils were pollinated with Yiduo pollens, and Yiduo pistils were pollinated with Shixia pollens; then, pistils were picked 0, 4, 8, 12, and 24 HAP separately. Isolated pistils were then fixed in Carnoy's fixative solution for 1 day. The pistils were then cut dorsally after gradient rehydration and softened with 4% NaOH for 12 h. The cut and softened pistils were stained with 0.1% aniline blue and pressed to desired thickness. Stigma and style were photographed under an upright fluorescence microscope (Zeiss AxioScope A1; Zeiss, Jena, Germany), and the ovule under a confocal laser scanning microscope (Zeiss LSM710; Zeiss, Jena, Germany) equipped with a Zeiss Axio Cam HRC camera (Zeiss, Jena, Germany).

Transcriptome Sequencing

Thirty samples (**Table 1**) were processed for extraction of total RNA, purification of mRNA, and quantification as reported earlier (Jue et al., 2019; Lee et al., 2019). The quantified mRNA samples were then used to construct libraries. For this purpose, the mRNA was enriched with magnetic beads [together with Oligo (dT)], and it was randomly interrupted by adding fragmentation buffer, first strand cDNA was synthesized. The purified double-stranded cDNA was repaired, A-tailed, and connected to the sequencing adapter, and then AMPure XP beads were used for fragment size selection. Finally, the cDNA library is obtained by PCR enrichment. The quality of the libraries was determined by quantitative PCR, the libraries were pooled for each sampled tissue (pollination combination), and sequencing was performed on an Illumina platform.

Bioinformatics Analyses of RNA-seq

Raw data were filtered to obtain clean data by removing reads containing connectors and low-quality reads in FastQC[2]. GC content distribution check was executed. The transcriptomic data were aligned with the longan reference genome (Lin et al., 2017) using HISAT2 (Kim et al., 2019). Comparison efficiency (percentage of mapped reads in clean reads) was computed and expressed as table and visualized in Integrative Genomics Viewer (Robinson et al., 2011).

Gene expression was quantified as fragments per kilobase of transcript per million fragments mapped (FPKM) and overall distribution of gene expression was expressed as a graph. Pearson's correlation coefficient (PCC) and principal component analysis (PCA) were computed for the expression data between replicates of the treatments in prcomp[3]. Gene (read) count was calculated in order to obtain false discovery rate (FDR). After this, differentially expressed genes (DEGs) were screened using a criterion, i.e., log2 fold change ≥ 2 and FDR < 0.01. DEG-related analyses were performed in DESq2 (Love et al., 2014).

The DEGs were functionally annotated in different databases, i.e., KEGG (Kanehisa et al., 2004), gene ontology (GO) (Ashburner et al., 2000), clusters of orthologous groups (COG) (Tatusov et al., 2000), PfAM, Swissprot (Apweiler et al., 2004), egNOG (Huerta-Cepas et al., 2016), NR (Deng et al., 2006), and KOG (Koonin et al., 2004) using BLAST (Altschul et al., 1997). We then performed enrichment of DEGs in KEGG pathways in R/clusterProfiler (Version 3.10.1).

RT-qPCR Analysis

To validate the RNA sequencing results, we selected 19 genes of interest (involved in JA metabolism) and performed RT-qPCR analyses. Primers were designed using Primer3Plus (Untergasser et al., 2007) (**Table 2**). qPCR setup and reaction conditions were

[2]http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[3]www.r-project.org

FIGURE 1 | Phenotype of the two reciprocal crosses between Yiduo (Y) and Shixia (S). **(A)** Yiduo♀ × Shixia♂(YS) showing no fruiting. **(B)** Shixia♀ × Yiduo♂(SY) showing fruiting.

TABLE 1 | Details of longan samples used for transcriptome sequencing.

	Time after pollination (hours)			
Pollination combination	4	8	12	24
Yiduo ♀ x Shixia ♂(YS)	YS4	YS8	YS12	YS24
Shixia ♀ x Yiduo ♂(SY)	SY4	SY8	SY12	SY24
CK1	Yiduo sepal-pistil before pollination			
CK2	Shixia sepal-pistil before pollination			

as reported earlier (Luo et al., 2021). Relative gene expression was computed using the *GAPDH* gene as an internal control (Luo et al., 2021).

Statistical Analyses

For biochemical (endogenous levels of the hormones) and phenotypic traits (time required by a tube to reach the end of style, length of style, and growth rate of pollen tube), means were compared in Microsoft Excel 2019 by one-way ANOVA. Difference at $p < 0.05$ was considered statistically significant. Standard deviation was also computed for each trait.

RESULTS

Pollen Germination and Pollen Tube Growth

The early datasets from our laboratory showed that in YS the fruit setting rate % is zero (**Figure 1**). To extend our observations, we explored some possibilities to which the UCI could be associated. First, we tested if the pollen of Shixia is viable/cannot germinate/sterile. To this regard, our microscopic observation showed that the pollens of both Yiduo and Shixia can germinate (**Supplementary Figure 1**). Second, we tested if the stigma of Yiduo is receptive or not. We found that the Yiduo stigma is receptive (**Supplementary Figure 1**), suggesting that the UCI is probably not due to defective stigma in Yiduo but another reason. Third, we tested if the pollen tube of Shixia can/cannot grow on the style of Yiduo. For this, we observed the *in situ* growth process of the pollen tube after pollinations in YS (**Figure 2I**). As expected, there was no growth of pollen tube prior to pollination (**Figures 2A–C**). In YS, the pollen tube elongated to the middle of the style 4 HAP (**Figures 2D–F**) and reached the end of the style 8 HAP (**Figures 2G–I**). However, 24 HAP, pollen tube signal became weaker (**Figures 2J–L**). This indicates that the pollen tube of Shixia can grow on the style of Yiduo. Finally, we observed that the pollen tube of Yiduo could enter into the ovule of Shixia, but in the case of Shixia pollen tube, it could not reach the ovule of Yiduo (**Figure 2II**). Altogether, the microscopic observations showed that the pollens of both Shixia and Yiduo can germinate, the stigmas of both cultivars are receptive, and the pollen tubes of both can growth into the style of the other. However, the final observation that Shixia pollen tube failed to reach Yiduo ovule is suspected to be the cause of the UCI (**Figures 2N,P**).

To further understand the possible cause of the failure of Shixia pollen tube to reach Yiduo ovule, we measured the lengths of styles of Yiduo and Shixia. We observed that the lengths of the styles of Yiduo and Shixia are different, i.e., the length of Yiduo style is almost double of that of Shixia style (**Figures 3A,B**). This suggests that the failure of Shixia pollen tube to reach Yiduo ovule (**Figure 2II**) can be linked to the length of Yiduo style (**Figures 3A,B**). Besides, we also measured pollen tube growth rate in the two reciprocal pollination combinations. We found that the growth rate of the pollen tubes was significantly higher in SY than in YS (**Figure 3C**). Altogether, our observations propose that the observed IUC in YS is caused by the failure of the pollen

TABLE 2 | List of primers used for real-time quantitative polymerase chain reaction (RT-qPCR) analyses of the selected genes.

Gene ID	Gene name	Forward primer sequence	Reverse primer sequence
Dlo_022603.1.gene	*PLA1*	ATCGTGGTCGTTGTCG	GAACTACAGAGGCGTGAG
Dlo_020248.1.gene	*PLA1*	ACGTTTGAGAGAA	AGAAATCACCCAGC
Dlo_014311.1.gene	*PLA1*	AGCTCCCACTACCT	TCAACTTCTGCCCCAC
Dlo_031198.1.gene	*LOX*	TTAGGCTATGGCGAGG	TAAGACGAGCA
Dlo_006067.1.gene	*LOX*	CCATGGTCAACCTCCTC	TCCACCTTTCATGTGCTC
Dlo_012185.1.gene	*LOX*	ACTCCGGTTAACAGC	TTACACACTACATTTC
dlo_037902.1.gene	*LOX*	GGTAGCAGGATCAATA	TGACCAAATGCAC
Dlo_012184.1.gene	*LOX13*	GAAGACATTTGAGA	ATTTAGGCCTAGTAT
Dlo_001507.1.gene	*LOX14*	TGTAGTGGGGTGGTAT	ATGTGGCATGAGGGG
Dlo_015582.1.gene	*AOS*	GATCGTGTTACATCA	TCTCAGCCTCAAAGT
Dlo_011584.1.gene	*AOC*	TGATTGGTGAAGCTCAA	TCATTGCACTGGC
Dlo_024807.1.gene	*OPR11*	CATCTGATGGCACA	TTACGAAGATGTGGG
Dlo_010077.1.gene	*OPR11*	ATTAGGGGCAATC	CCAACTCTTCGAAGTG
Dlo_022217.1.gene	*OPR11*	GTTGCAAGTTCCATACT	CAGCATAGCCAAGTCG
Dlo_001987.1.gene	*JMT*	ATCCTAAGGGATGC	AGGATATTTCGCTG
Dlo_012262.1.gene	*JAR1*	GCCAGTCATTCAAGCA	GCAGGTTCAAGCTTA
Dlo_003095.1.gene	*CYP94A2*	TCCACCCTATGCTCTCA	AGCCTCAATGGCAA
Dlo_013441.1.gene	*CYP94C1*	AGGCAGATATTCATAG	CTCAGAGCAGATAG
Dlo_032545.2.gene	*CYP94C1*	GTCCACATACTTCACG	AGAGAACAGATGAT
GAPDH		AACGTTGCCTGATTTT	GTACTTTCTTTCATACT

tube of Shixia to enter the ovule of Yiduo, and this is probably due to the slow growth rate of the pollen tube of Shixia or the length of the style of Yiduo, or a combination of both phenomena.

Transcriptome Sequencing

Our phenotypic and microscopic observations indicated chances of involvement of pollen tube growth and failure of pollen to reach the ovule in the case of YS pollination combination. Therefore, we further explored the transcriptomic signatures of both pollination combinations 4, 8, 12, and 24 HAP and compared them with the respective non-pollinated Yiduo (CK1) and Shixia (CK2) sepal-pistil.

Transcriptome sequencing of the 30 libraries (**Table 1**) resulted in 271.9 GB clean data; on average, we got 5.71 Gb per sample with a Q30 base percentage of >93.7%. Comparison efficiency with the reference genome was 89.11 to 94%, and GC content was higher than 44%. On average, >90% of the reads could be mapped to the reference genome (**Supplementary Table 1**). Overall gene expression was lower in CK than in the other treatments (**Figure 4A**). The PCA showed that all biological replicates clustered together. In addition, all the treatments were grouped close to each other except for CK1, CK2, and SY24 (**Figure 4B**).

Differential Gene Expression

Screening conditions for the DEGs were fold change $\geq$ 2 and FDR $<$ 0.01. These conditions resulted in the identification of 7,251 DEGs in the studied longan pollination combinations. Overall, we observed that the number of downregulated DEGs was higher than that of upregulated DEGs in all the pollination combinations compared to their respective controls (**Figure 4C**). Relatively higher number of genes were differentially regulated between CK1 and YS time points indicating large scale transcriptomic regulation in YS as compared to SY in all time points as compared to their respective controls.

The KEGG pathway enrichment analysis showed that the DEGs were enriched in plant-hormone signal transduction pathway (93 DEGs), protein processing and endoplasmic reticulum (70 DEGs), and starch and sucrose metabolism (81 DEGs). Other than these, we also found enrichment of DEGs in phenylpropanoid biosynthesis pathway (66 DEGs). All these pathways were common in all the treatment comparisons (**Supplementary Figure 2** and **Supplementary Table 2**).

First of all, we compared the transcriptome sequencing results of both SY and YS (4, 8, 12, and 24 HAP) with their respective controls to find out DEGs associated with pollination. There were 769 DEGs common in all the time points in YS compared to CK1. Contrastingly, there were only 171 DEGs common in SY compared to CK2 (**Figure 5** and **Supplementary Table 3**). These common DEGs could be related to overall development of the studied tissues in each of the pollination combinations. We found that 778 and 229 genes were specifically regulated in CK1 vs. SY24 and CK2 vs. YS24, respectively. This is the time point when we expect that the pollen will reach the ovule in the case of SY. The upregulated genes included ABC transporter G family members, ankyrin repeat-containing protein, *bHLH131*, *bHLH25*, and *WRKY33* TFs in SY24 compared to CK2. Interestingly, we also found the upregulation of an anther-specific protein, *LAT52,* in SY24 compared to CK2. This gene has been characterized for its roles in pollen development; however, its specific expression in SY24 might indicate its role in pollen tube development (Sheoran et al., 2007). Other than this, there were two genes, i.e., an ABC transporter and a later embryogenesis abundant

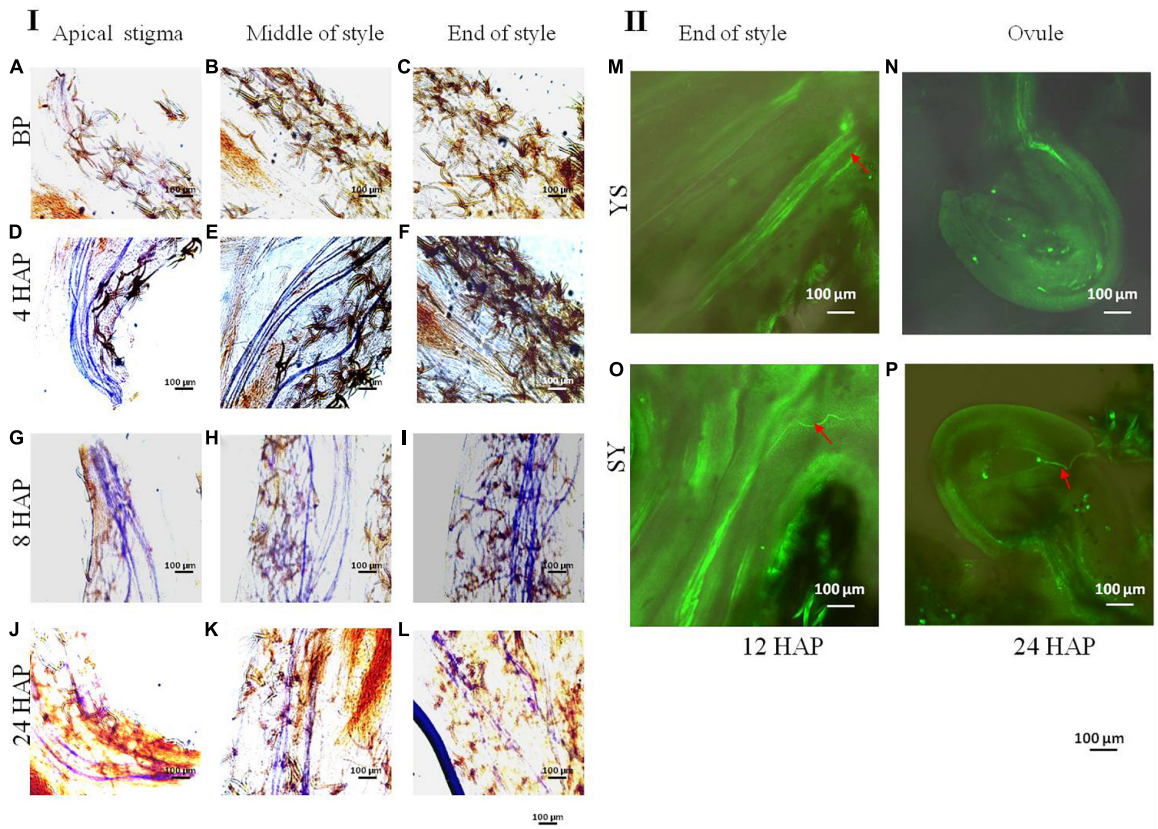


FIGURE 2 | Observation of dynamic growth process of pollen tube. **(I)** Observation of dynamic growth process of pollen tube before and after pollination with Shixia pollen by *in situ* germination under an upright fluorescence microscope. **(A)** Stigma before pollination. **(B)** Middle of style before pollination. **(C)** End of style before pollination. **(D)** Stigma 4 h after pollination. **(E)** Middle part of style 4 h after pollination. **(F)** End part of style 4 h after pollination. **(G)** Stigma 8 h after pollination. **(H)** Middle part of style 8 h after pollination. **(I)** End part of style 8 h after pollination. **(J)** Stigma 8 h after pollination. **(K)** Middle part of style 8 h after pollination. **(L)** End of style 24 h after pollination. The pollen tube is blue after being stained with ultraviolet light. **(II)** Observation of the dynamic growth process of pollen tube by *in situ* germination method under a confocal laser scanning microscope in reciprocal crosses of Yiduo and Shixia. **(M)** Yiduo × Shixia 12 h after pollination. **(N)** Yiduo × Shixia 24 h after pollination. **(O)** Shixia × Yiduo 12 h after pollination. **(P)** Shixia × Yiduo 24 h after pollination. The red arrow points to the pollen tube.

protein (*Dimocarpus_longan_newGene_11049*), that were highly expressed in SY24 compared to CK2.

Differential Regulation of Plant-Hormone Signal Transduction Pathway

All the treatment comparisons showed KEGG pathway enrichment of plant-hormone signal transduction, suggesting an important role of this pathway in pollination. Therefore, we focused on the DEGs enriched in this pathway. Ninety-three DEGs were enriched in the plant-hormone signal transduction pathway (**Supplementary Table 4**; see DEGs highlighted in dark orange color; **Figure 6**). We observed that almost all the phytohormone signaling-related pathways were differentially regulated in at least one time point in either of the pollination combination. Interestingly, we found the specific regulation of 11 DEGs only in SY (DEGs with * in the **Figure 6**). These included BAK1, two ERF1s, GID1, IAA4, JAZ, PRP1, SAUR, SRK2, and two TGAs. On the contrary, a larger number of DEGs (42) were specifically expressed in CK1 vs. YS (4, 8, 12, and 24 HAP). These observations indicate that phytohormone signaling has specific roles in UCI in the pollination combination YS.

Differential Expression of Jasmonic Acid Metabolism

As reviewed in the section "Introduction," JA levels have been related to limited pollen germination in Arabidopsis (Ju et al., 2016). Therefore, we explore JA metabolism to see the differential transcriptome signature in both pollination combinations. The conversion of galactolipids into α-linoleic acid by PLA1 and DAD1 initiates JA metabolism in plants (Ishiguro et al., 2001). Although we did not observe the differential expression of DAD1 gene between the studied samples, we observed that two PLA1 genes (*Dlo_020248.1.gene* and *Dlo_022603.1.gene*) had relatively lower expression in the YS series compared to the CK1 and SY time points. Particularly, their expression was nearly half in YS compared to SY (**Figure 7**). In addition, we observed that six transcripts annotated as LOXs were differentially regulated in the studied tissues. Of these, one LOX (*Dlo_012185.1.gene*) was specific to YS, and another was specific to SY (*Dlo_012184.1.gene* LOX13), while the others were common in both pollination

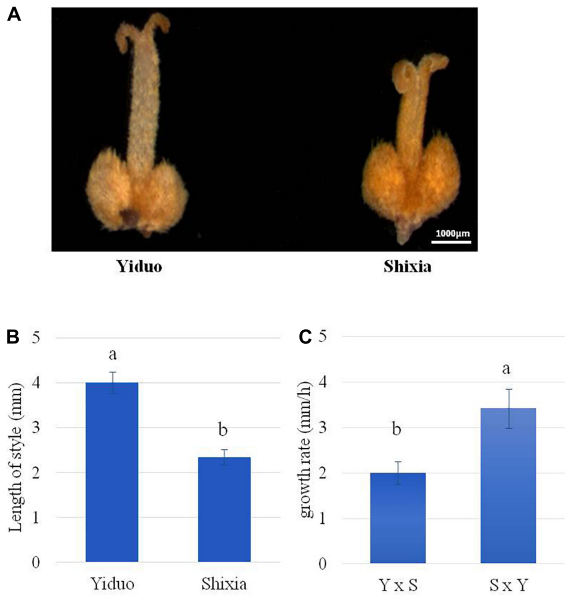


FIGURE 3 | Comparison of the pistil and pollen tube of longan with a stereomicroscope. **(A)** Observation of the pistil of Yiduo and Shixia with a stereomicroscope. **(B)** Length of the style in mm and **(C)** comparison of pollen tube growth rate in the two reciprocal crosses. Y, Yiduo; S, Shixia. Different letters show that the means differ significantly ($p < 0.05$). Error bars represent the standard deviation.

combinations. The expression trend of the LOX genes was similar in both YS and SY pollination time series. However, two LOXs (*Dlo_006067.1.gene* and *Dlo_012185.1.gene*) had globally higher FPKM values in YS than in SY, suggesting that 24 HAP, the level of JA is higher in YS than in SY (**Figure 7**). AOS and AOC control the conversion of 13(s) hydroperoxylinoleic acid to 12-13-epoxylinoleic acid to 12-oxophytodienoate (12-OPDA), respectively. 12-OPDA is then reduced to OPC8 (3-oxo-2-(2-pentenyl)-cyclopentane-1-octanoic acid). This reaction is controlled by OPR genes (OPDA reductase) (Wasternack and Hause, 2013) and finally leads to JA biosynthesis. In this study, we identified one AOS gene with a conspicuous different expression pattern between YS and SY. We observed that the transcript annotated as AOS (*Dlo_015582.1.gene*) had globally higher expression in YS than in CK1. However, in SY, its expression was globally lower than in CK2. The transcriptome analysis also showed three transcripts annotated as OPR11 (*Dlo_010077.1.gene, Dlo_022217.1.gene*, and *Dlo_024807.1.gene*) that were differentially expressed in the studied tissues. Only *Dlo_010077.1.gene* showed an obvious difference between YS and SY. The comparison between YS24 and SY24 indicated a ∼three-fold difference (YS24 had higher expression), suggesting high conversion of 12-OPDA to OPC8 in YS. Collectively, higher expression of JA biosynthesis genes in YS than in SY was observed, which implies high JA level in YS tissues. JA is converted into MeJA by JA carboxyl methyltransferase (JMT) (Seo et al., 2001). The transcriptome sequencing showed that a JMT gene (*Dlo_001987.1.gene*) had reduced expression in YS4, YS8, and YS12 compared to CK1, but that its expression significantly increased in YS24 compared to CK1. In the case of SY, its expression decreased in all the four time points compared to CK2; however, the expression in SY24 was relatively higher 4, 8, and 12 HAP. The expression in both combinations was quite similar 24 HAP, proposing a similar concentration between the two combinations 24 HAP. Overall, the comparative expression analysis directs that the levels of JA might be higher in the YS time points than those in the SY pollination combination (the JA content in the studied comparison is presented below in section "Endogenous Jasmonic Acid Levels in Pistils of Yiduo and Shixia") (**Figures 6**, **7**).

Differential Expression of Jasmonic Acid Signaling

Since JA signaling has been shown to be involved in the regulation of genes associated with pollen germination, pollen tube growth, and cross incompatibility (Muradoğlu et al., 2010; Guan et al., 2013; Chaturvedi et al., 2016), we explored the changes in the expression of genes that were involved in JA biosynthesis and signaling. JA signaling is a part of the plant-hormone signal transduction pathway (Yang et al., 2019). JA-Ile is formed by jasmonoyl-isoleucine synthetase 1 (JAR1), which is a member of the GH3 gene family (Staswick and Tiryaki, 2004). This is the major step in JA perception, since JA-Ile is the most biologically active JA compound (Fonseca et al., 2009). The expression of JAR1 was lower in YS4 and YS8 than in CK1, but it was increased in YS12 and YS24; in particular, the expression was significantly higher 24 HAP, suggesting higher conversion of JA into JA-Ile at this time point. A similar expression pattern was noted in the four time points of the SY pollination combination. The main difference between both pollination combinations was that YS24 had higher FPKM values than SY24, proposing that YS24 might have higher JA-Ile concentration than SY24. JA-Ile signal is perceived by coronatine-insensitive protein 1 (COI), which triggers the repression of JAZ. In both pollination combinations, the expression of JAZs decreased as compared to their respective controls, indicating that JA-Ile triggered the repression of JAZs, although we did not note the differential regulation of COI. JAZ repression releases downstream transcription factors (TFs), e.g., MYC2s. We noted that two MYC2s were differentially regulated between the two pollination combinations. Both MYC2 genes (Dlo_016741.1.gene and Dlo_027394.1.gene) were downregulated in YS as compared to CK1, whereas only downregulated in SY4 and SY8 but upregulated SY12 and SY24 as compared to CK2. This observation proposes that in the YS pollination combination, the JA-responsive activation of MYC2 is relatively lower than that of SY (in particular 12 and 24 HAP). There are multiple genes present downstream of MYC2s that regulated JA-triggered responses in plant tissues (Ruan et al., 2019) (**Figures 6**, **7**). Overall, considering the regulation of JA metabolism and signaling-related genes in the two pollination combinations, it can be proposed that JA biosynthesis and signaling have an important role in UCI in YS.

The expression pattern of 19 genes involved in JA metabolism was also confirmed by RT-qPCR analysis (**Supplementary Figure 3**), which was consistent with the FPKM values of these genes. Theses expression profiles signify both the reliability of the RNA sequencing and the role of JA metabolism.

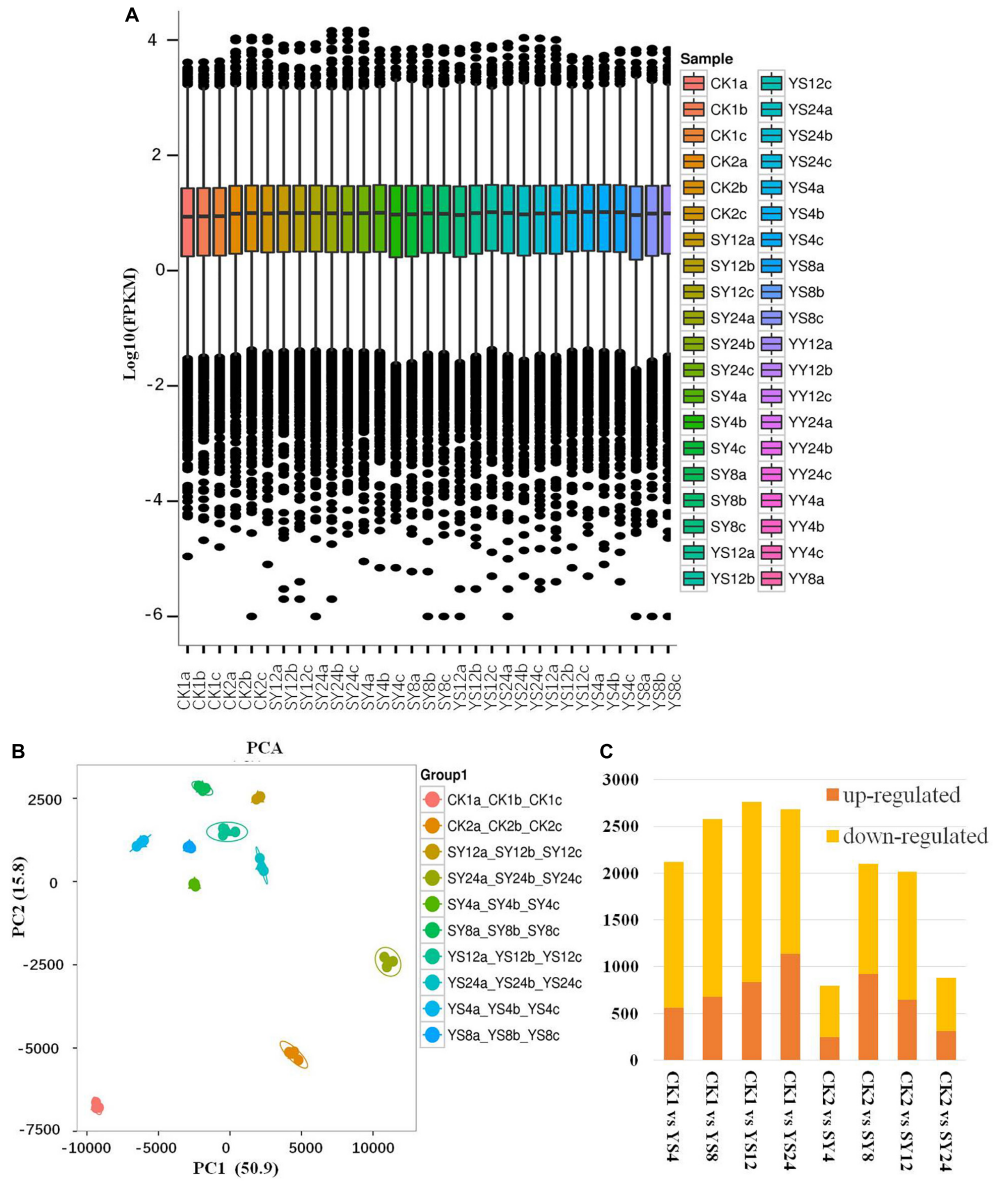


FIGURE 4 | Transcriptional data analysis of Yiduo and Shixia. **(A)** Overall distribution of gene expression. **(B)** Principal component analysis of the gene expression data among the replicates of the treatments. **(C)** Summary of the differentially expressed genes in longan sepal pistils. Y, Yiduo; S, Shixia; CK1, Yiduo sepal-pistil before pollination, and CK2, Shixia sepal-pistil before pollination. The numbers with treatments represent time after pollination.

Differential Regulation of Phenylpropanoid Pathway, Cell Wall, and Pollen Tube-Related Genes

The KEGG pathway analysis indicated that 66 DEGs were enriched in the phenylpropanoid pathway. The DEGs included multiple genes that control the main steps of the phenylpropanoid biosynthesis pathway. Three β-glucosidase (BGLs), a caffeic acid (3-O-methyltransferase, CA3OM), two CADs, two shikimate O-hydroxycinnamoyltransferase (HSTs), a phenylalanine ammonia-lyase (PAL), and seven peroxidases (PODs) were specific to SY 8 and 12 HAP. Similarly, 21 DEGs were specific to YS. We observed that of the three 4CLs, the expression of two transcripts (*Dlo_021761.1.gene* and *Dlo_024529.1.gene*) was significantly decreased in all the YS time points, whereas the expression of the third transcript increased significantly in YS24 compared to CK1 and YS4-12. Contrastingly, the expression of the three 4CL transcripts decreased at SY4 and SY8 but then returned to nearly the same as of CK2 (but remained lower than CK2). One CA3OM transcript showed an expression pattern similar to that of 4CLs. The biosynthesis of coumarine might be similar in both pollination combinations because of increase in the expression of BGLs. BGLs are known for their roles in the biosynthesis of coumarins (Schaeffer et al., 1960). How coumarin biosynthesis may/may not affect pollen tube elongation is less known (Adhikari et al., 2020).

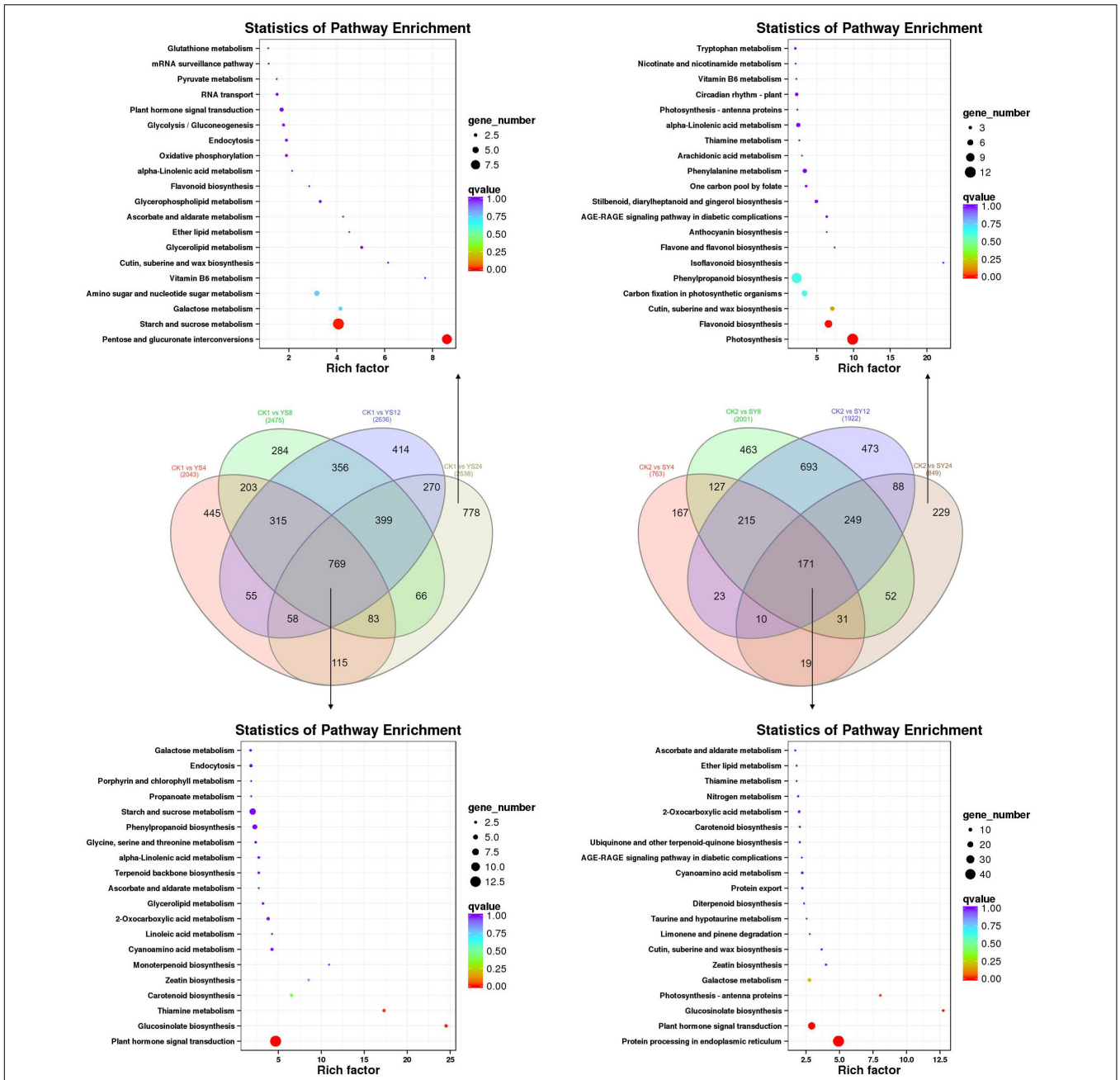


FIGURE 5 | Differentially expressed genes in pollination combinations YS and SY. Venn diagrams show the common and specific DEGs in each pollination combination. KEGG enrichment scatter plots indicate the pathways in which common and specific DEGs were enriched. Y, Yiduo; S, Shixia; CK1, Yiduo sepal-pistil before pollination; CK2, Shixia sepal-pistil before pollination. The numbers with treatments represent time after pollination.

However, current observations propose that coumarins have a limited role in pollen tube elongation. We also found the differential expression of PALs, where two of the three PAL transcripts showed reduced expression compared to the controls, while the third transcript showed increased expression. This similar expression trend between both pollination combinations might suggest that PAL might not be a key stage for the control of the biosynthesis of lignin in longan. This is understandable, since PAL controls the first stage of phenylpropanoid biosynthesis, i.e., phenylalanine conversion into cinnamic acid (Zhang and Liu, 2015) (**Supplementary Table 4** and **Figure 8**). Our transcriptome sequencing also showed the differential expression of a transcript annotated as myo-inositol transporter 13 (INT13). Its expression decreased in both YS and SY series compared to CK1 and CK2, respectively; however, its expression in YS24 was lower than in SY24.

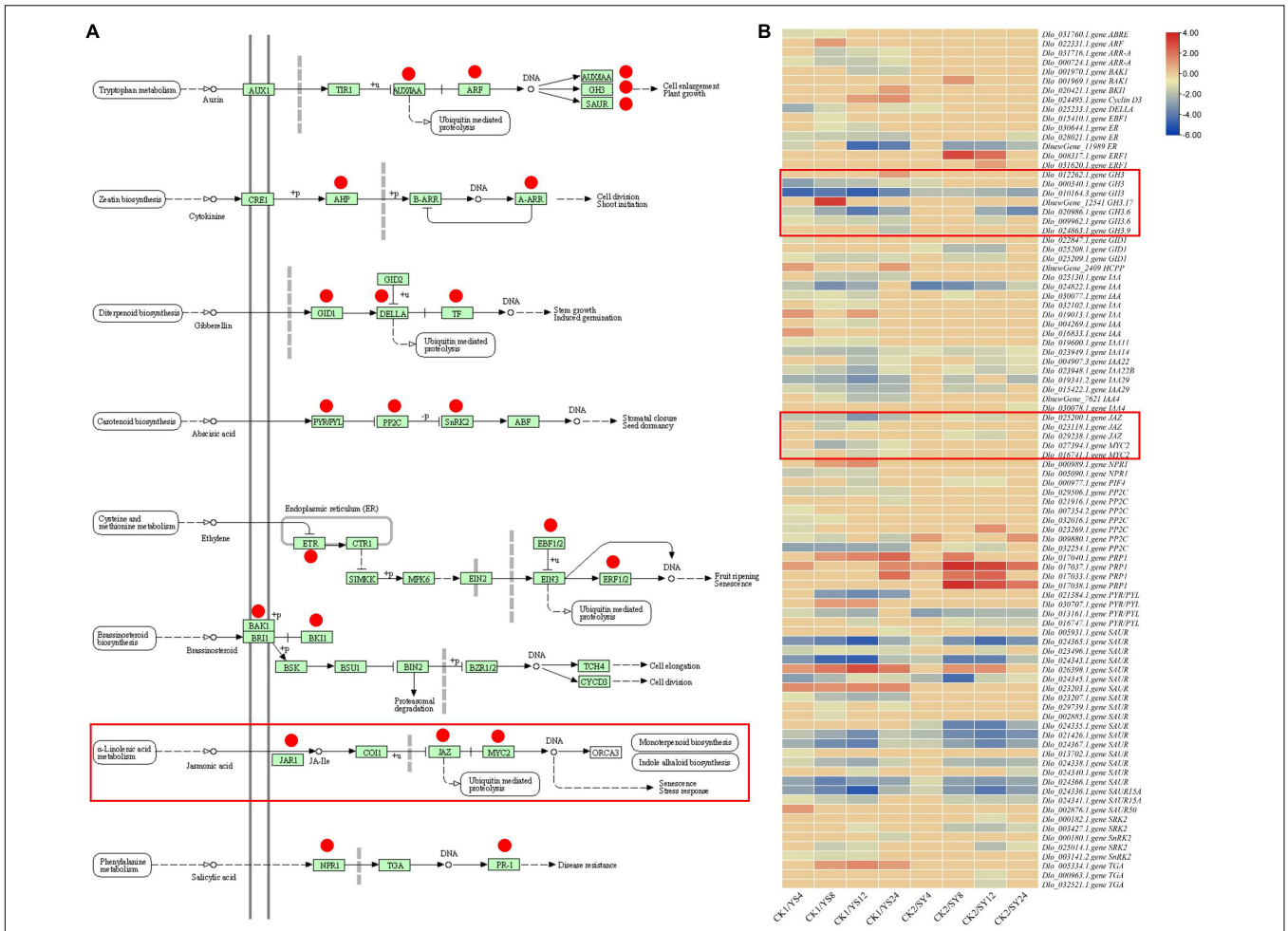


FIGURE 6 | Comparison of hormone signal transduction pathway in the two longan pollination combinations. **(A)** Regulation of plant-hormone signal transduction pathway in two longan pollination combinations, i.e., YS and SY, compared to their respective controls. Y, Yiduo; S, Shixia; CK1, Yiduo sepal-pistil before pollination; CK2, Shixia sepal-pistil before pollination. The numbers with treatments represent the time (hour) after pollination. The red circles in the left panel represent the genes that were differentially regulated. Red rectangles represent the genes related to Jasmonic acid. **(B)** The right panel shows the heatmap of log2 fold change values of the DEGs between YS (4, 8, 12, and 24 HAP) vs. CK1 and SY (4, 8, 12, and 24 HAP) vs. CK2.

Since, pollen tube is a fast-growing cell and requires higher amounts of cell wall deposition, it is also essential to check the expression of cell wall-related genes. We selected all the DEGs (33) containing the term "cell-wall" based on GO annotation and checked their differential regulation in SY and YS. Of the 33 cell wall-related DEGs, 15 were specific to YS. We found that the expression of a cellulose synthase A (CESA, *Di_newGene_10844*), galacturan 1,4-alpha-galacturonidase (GAG, *Dlo_021956.1.gene*), and a pectinesterase (PE, *Dlo_010992.1.gene*) was reduced in SY (4, 8, 12, and 24 HAP) compared to CK1. These expression changes may suggest that the deposition of cell wall in the growing pollen tube might be lower in the case of YS. However, pollen tube growth in both SY and YS, at least until 12 HAP, could be due to the higher expression of a chitinase (CHI, *Dlo_030519.1.gene*) and a CESA (*Dlo_017102.2.gene*). Particularly, we found that *Dimocarpus_longan_newGene_10844* (CESA) expressions were reduced compared to CK1 in all the four time points. Furthermore, its expression remained almost unchanged in SY4-8, slightly decreased in SY12, and then increased in SY24 (although in this stage, the expression was still lower than in CK2). In YS24, its expression was lower than that in SY24. Other than CESAs, we found that the expression of most of the xyloglucan:xyloglucosyl transferases (XETs) was higher in YS24 than in CK1 and at all the SY time points. Similarly, we noted the differential expression of xylanase inhibitors (XIPs, *Dlo_001017.1.gene* and *Dlo_022672.1.gene*), pectin methylesterase (PME, also known as PE, *Dlo_023530.1.gene*), and PME-inhibitor (PMEI, *Dlo_024765.1.gene*).

We also searched for the term "pollen tube" to screen the DEGs associated with pollen tube and found 22 genes that were differentially expressed in CK1 vs. YS and CK2 vs. SY at different time points after pollination. We noted that a cation/H(+) antiporter 15 (CHX15, *Dlo_011759.1.gene*)

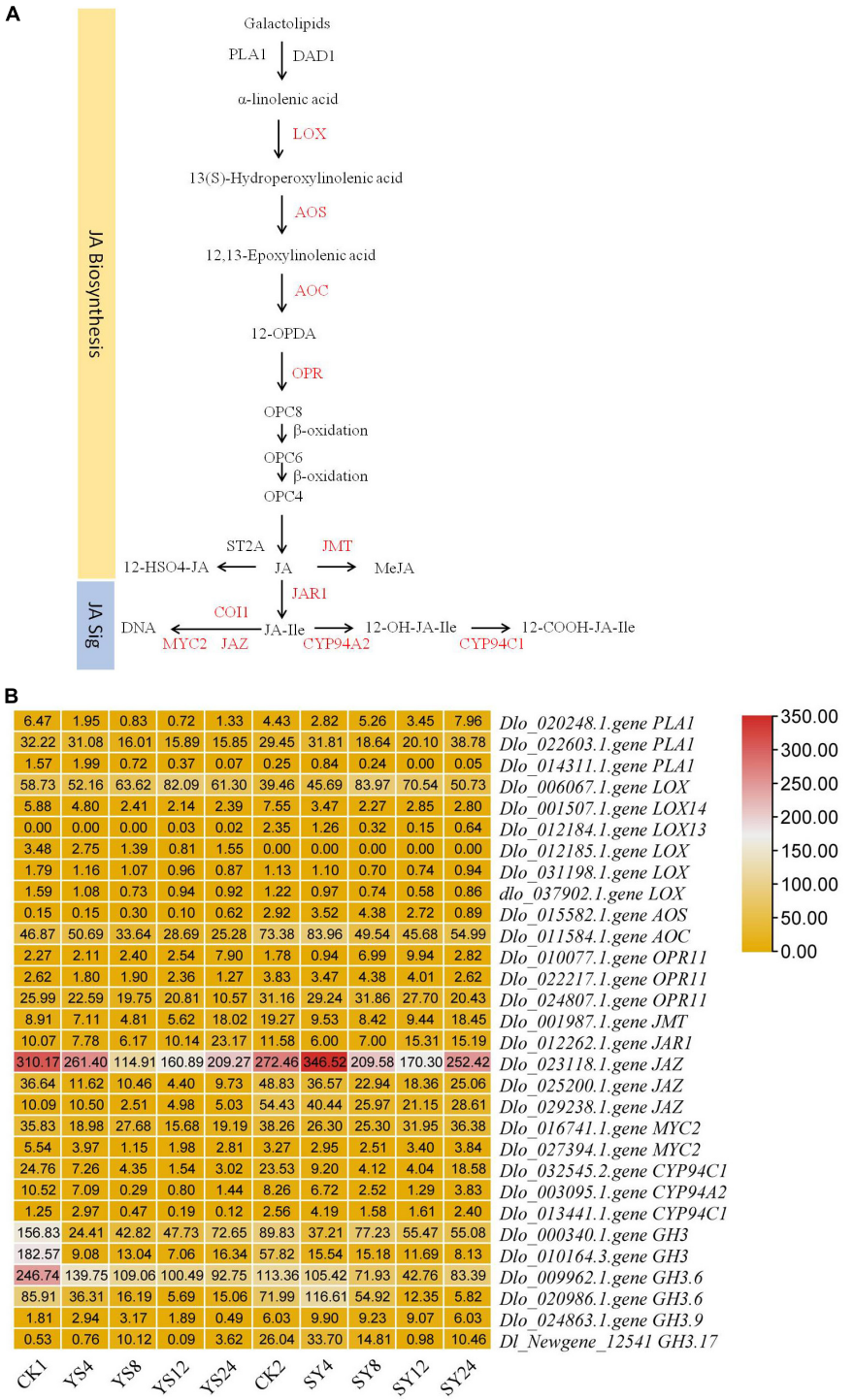


FIGURE 7 | Differential regulation of jasmonic acid metabolism and signaling in the YS and SY pollination combinations compared to their respective controls. **(A)** The pathway of Jasmonic acid metabolism and signaling. The red colored text represents the differentially regulated genes in the pathway. DAD1 (DEFECTIVE IN ANTHER DEHISCENCE 1), LOX (lipoxygenase), AOS (allene oxide synthase), AOC (allene oxide cyclase), OPR (OPDA Reductase), JMT (JA methyl transferase), ST2A (sulfotransferase 2A), JAR1 (Auxin-responsive GH3 family protein; jasmonic acid-amino synthetase), COI1 (coronatine-insensitive protein 1), JAZ (Jasmonate ZIM domain-containing protein), MYC2 (transcription factor MYC2), CYP94A2 (jasmonoyl-isoleucine 12-hydroxylase), and CYP94C1. Y, Yiduo; S, Shixia; CK1, Yiduo sepal-pistil before pollination; CK2, Shixia sepal-pistil before pollination. The numbers with treatments represent time after pollination. **(B)** The heatmap shows the FPKM values of differentially regulated genes.

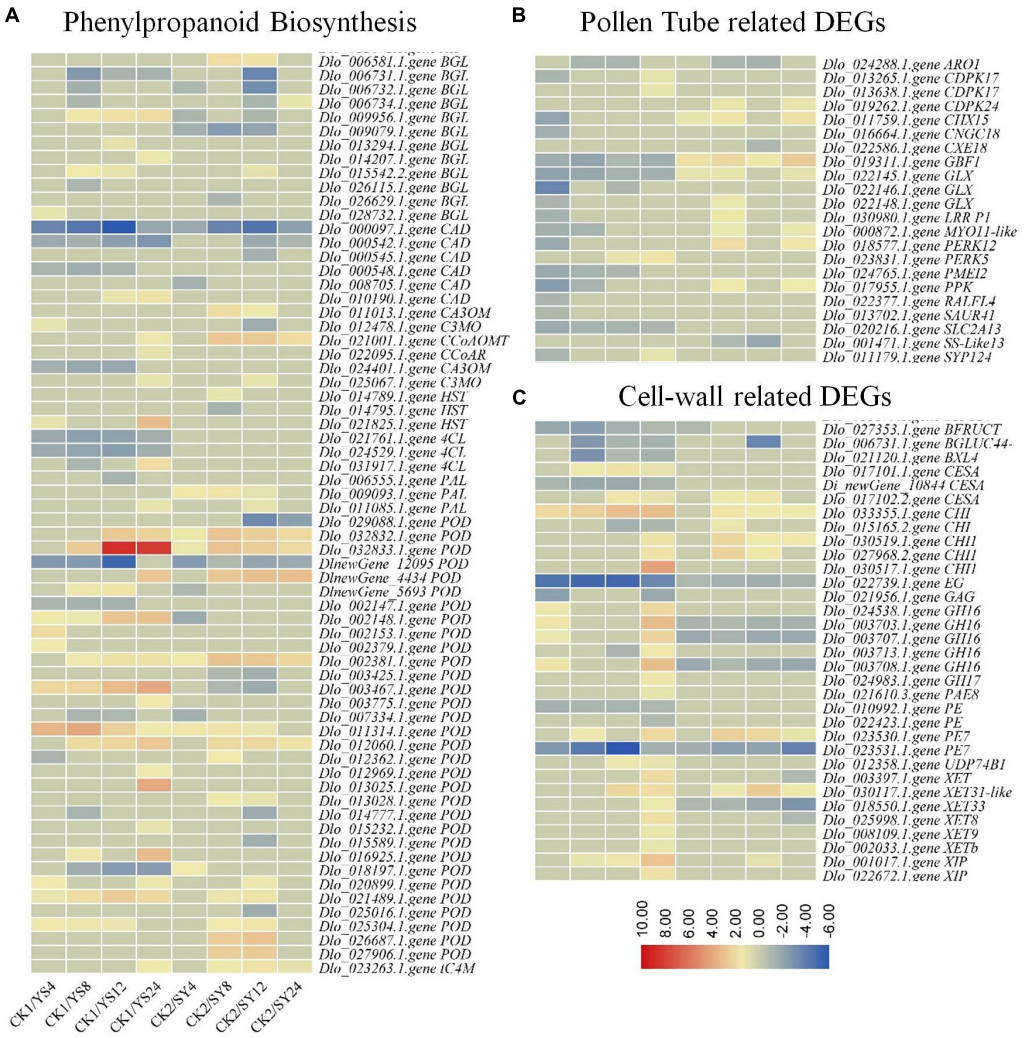


FIGURE 8 | (A) Differential regulation of phenylpropanoid biosynthesis pathway, **(B)** pollen tube-related, and **(C)** cell wall-related genes in the CK1 vs. YS series and CK2 vs. SY series. The heatmaps show log2 fold change values of the DEGs. Y, Yiduo; S, Shixia; CK1, Yiduo sepal-pistil before pollination; CK2, Shixia sepal-pistil before pollination. The numbers with treatments represent time after pollination.

showed higher expression in SY24 than in CK2, whereas its expression first reduced 4,8, and 12HAP in YS but then slightly increased (although not differentially expressed) in YS24 compared to CK1. Regarding our search against the term "pollen tube," we found the differential expression of many Ca^{2+} signaling-related genes. Particularly, we found the differential expression of three calcium-dependent protein kinase (CPK) genes (CPK24 and two CPK17), a cyclic nucleotide-gated channel (CNGC18), and a syntaxin-124 (SYP124) gene. CPK17s were highly expressed in the YS time points (particularly, their expression increased in YS24 compared to CK1), whereas their expression was only fractional in the SY time series. CNGC18 (*Dlo_016664.1.gene*) showed higher gene expression in YS24 than in SY24. The transcriptome data also revealed the higher expression of SYP124 in YS than in SY. This expression trend is similar to that of other genes affected by Ca^{2+} signaling. Other than these, we noted the differential expression of a protein RALF-like 4 (RLAFL4, *Dlo_022377.1.gene*), a pollen-specific leucine-rich repeat extensin-like protein 1, STRICTOSIDINE SYNTHASE-LIKE 13 (SS-like 13, *Dlo_001471.1.gene*), an armadillo repeat only 1 (ARO1, *Dlo_024288.1.gene*), and two proline-rich receptor-like protein kinases (PERK12, *Dlo_018577.1.gene* and PERK5, *Dlo_023831.1.gene*). The RALF protein is known for male sterility (Zou et al., 2017), whereas PERKs are known for pollen tube growth (Borassi et al., 2021). RLAFL4, SS-like 13, and PERKS should be given due consideration in future studies exploring UCI in longan based on the expression patterns observed in our study.

Differential Regulation of Starch and Sucrose Biosynthesis Pathway

Earlier reports have presented that sucrose concentration affects significantly pollen tube growth (Fragallah et al., 2019).

Considering that we observed changes in cell wall biosynthesis, loosening, and remodeling-related genes (see section "Differential Regulation of Phenylpropanoid Pathway, Cell Wall, and Pollen Tube-Related Genes"), in addition, we checked the differential expression of genes (50 DEGs) that were significantly enriched in the starch and sucrose biosynthesis pathway. Interestingly, we found that three of the four DEGs annotated as sucrose synthases (SuSy, *Dimocarpus_longan_newGene_4925*, *Dlo_005657.1.gene*, and *Dlo_013696.1.gene*) were downregulated in YS compared to CK1. On the contrary, two of the four transcripts showed upregulation in SY, *Dimocarpus_longan_newGene_4925* and *Dimocarpus_longan_newGene_11275*), whereas one gene (*Dlo_005657.1.gene*) had stable expression in the SY series compared to CK2. Two sucrose phosphate synthases (SPSs, *Dlo_015436.1.gene* and *Dlo_022163.1.gene*) also showed decreased expression in YS compared to CK1, in contrast to the upregulation of the second transcript in SY compared to CK2 (**Supplementary Table 4**). These observations are consistent with the observation that in YS, pollen tube growth rate was lower than in SY (**Figure 3C**).

The 1,4-alpha-glucan branching enzyme (GBE, *Dlo_028752.1.gene*) gene was upregulated in all the four time points of both pollination combinations compared to their respective CKs. Two hexokinases (HKs, *Dlo_000488.1.gene* and *Dlo_027870.1.gene*) were downregulated in the YS series compared to CK1, whereas in SY, the first transcript was upregulated. Another enzyme fructokinase (FRK, *Dlo_022066.1.gene* and *Dlo_033973.1.gene*) that controls the same step as of HK i.e., conversion of D-fructose to D-fructose-6P (Jiang et al., 2003; Kawai et al., 2005), was downregulated in YS as compared to CK1 and upregulated in SY as compared to CK2. However, β-fructofuranosidase (BFRUCTF) showed decreased expression in both pollination combinations, indicating depletion of sucrose (Schwebel-Dugué et al., 1994). We also noted the differential expression of multiple genes that take part in D-glucose biosynthesis in the starch and sucrose biosynthesis pathway. The breakdown of cellulose to cellodextin by endoglucanase (EG, *Dlo_022739.1.gene* and *Dlo_028821.1.gene*) (Opassiri et al., 2006) and then to cellobiose by EG and beta-glucosidases (BGLs) (Yoshida and Komae, 2006) leads to D-glucose biosynthesis. The expression of EGs decreased in both YS and SY compared to CK1 and CK2, whereas the expression of BGLs was variable in both cases (**Supplementary Table 4**).

Finally, there were changes in the expression of genes related to starch metabolism. Glucose-1-phosphate adenylyltransferase (G1P), which converts α-D-glucose-1P to ADP-glucose (Ghosh et al., 1992), showed a similar expression pattern in both pollination combinations. Similarly, the gene that controls the biosynthesis of starch/glycogen from amylose i.e., 1,4-alpha-glucan branching enzyme (Baecker et al., 1986) was upregulated in all YS series as compared to CK1. However, this gene was downregulated in SY series (except for SY24). These expression changes indicate that the starch biosynthesis is higher in all YS time points but not in SY series (except for SY24) as compared to CK1 and CK2. Our transcriptome data showed that the expression of α-amylase (AAs) increased in the YS series compared to CK1. On the contrary, only one AA (*Dlo_001166.1.gene*) was upregulated in the SY series. The expression of two of the three β-amylases (BAs) (*Dlo_026796.1.gene* and *Dlo_029595.1.gene*) was reduced in YS compared to CK1 in all the time points. On the contrary, the expression of the two above-mentioned BAs was reduced in SY4-12 but significantly increased in SY24 compared to CK2 and CK1 (**Supplementary Table 4**).

Differential Regulation of Protein Processing in the Endoplasmic Reticulum

Pollen tube growth needs delivery of a wide variety of proteins and requires intense streaming and trafficking (Scali et al., 2021). In this regard, it is essential for a growing cell to process proteins in the endoplasmic reticulum. The KEGG pathway enrichment analysis indicated that 70 DEGs were enriched in protein processing in the endoplasmic reticulum pathway in both pollination combinations. Of the 70 DEGs, 40 were specifically expressed in CK1 vs. YS. Among the DEGs, two (HSP70KDa 1/8, *Dlo_032043.1.gene* and molecular chaperon HtpG, *Dlo_033047.1.gene*) showed interesting expression profiles, i.e., they were downregulated in all the treatment comparisons except in CK2 vs. SY24. Hsp70KDas are known for their involvement in polar nuclei fusion during female gametophyte and sperm nuclear fusion with central cell polar nuclei at fertilization (Maruyama et al., 2010). Also, they are required for pollen development and pollen tube growth (Maruyama et al., 2014). On the other hand, the role of HtpGs in pollen tube development is not elaborated yet. Nevertheless, HtpGs have been implicated in heat stress endurance of pollen tubes (Chaturvedi et al., 2016) (**Supplementary Table 4**).

Endogenous Jasmonic Acid Levels in Pistils of Yiduo and Shixia

Using ab SCIEX QTRAP 6500 LC-MS/MS technique, we measured the changes in JA and JA-Ile contents in the two pollination combinations, i.e., SY and YS, before and after 4, 8, 12, and 24 HAP (**Figure 9**). We observed that, overall, the JA levels were higher in YS than in SY (**Figure 9A**). A similar trend was noticed for the endogenous JA-Ile levels (**Figure 9B**). These observations propose that JA might be involved in UCI.

DISCUSSION

Cell Wall-Related Genes That Are Associated With Changes in Pollen Tube Growth Are Differentially Expressed in Shixia♀ × Yiduo♂ and Yiduo♀ × Shixia♂

There are many requirements for successful fertilization in plants, i.e., pollen germination, receptivity of the stigma, pollen tube

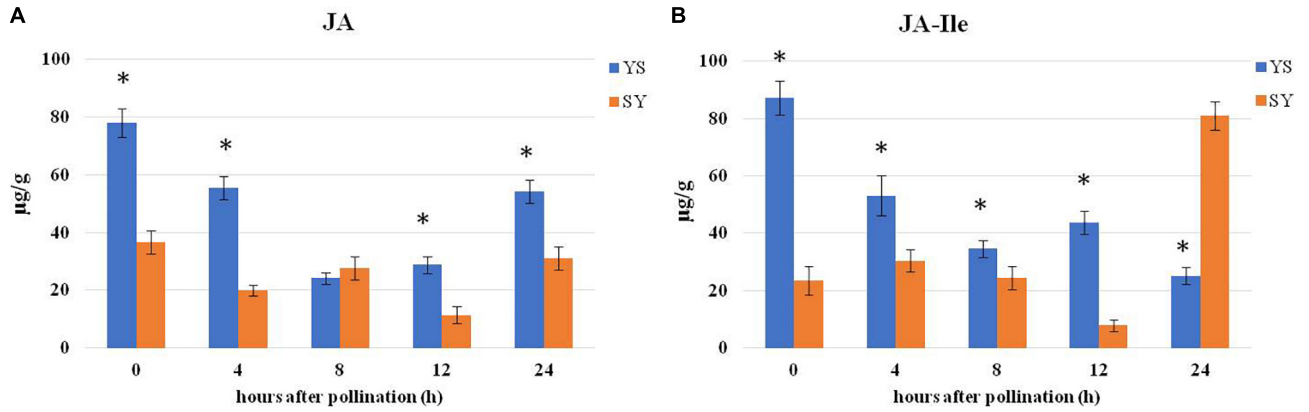


FIGURE 9 | Comparison of hormone levels in pistils of SY and YS. **(A)** Jasmonic acid content. **(B)** JA-Ile content 0, 4, 8, 12, and 24 h after pollination. YS, Yiduo♂ × Shixia♀; SY, Shixia♀ × Yiduo♂. Asterisk on the bars show significant differences ($p < 0.05$).

growth, and successful transmission of pollen (Krichevsky et al., 2007). Our observations that the pollens of both Shixia and Yiduo germinated normally, and that the stigma of the Yiduo was receptive (**Supplementary Figure 1**) show that these two pre-fertilization reproductive barriers are not different in the studied pollination combinations, SY and YS (Deng et al., 2017). The microscopic observation of the dynamic process of pollen tube growth indicated that when a Shixia pollen is used to fertilize a Yiduo stigma, signals of pollen tube growth weaken 24 HAP. This was further confirmed by observing the arrival of the pollen tube in the ovule 12 and 24 HAP. These results indicated that in the case of SY, the pollen tube reached the ovule within 24 HAP; however, in the case of YS, its growth rate was slower (**Figure 2II**). These microscopy-based observations propose that (1) stigmas of both S and Y are receptive; (2) in the case of SY, the pollen tube elongates and reaches the end of style in 24 h; (3) whereas in the case of YS, the pollen tube elongates into the style by 24 h but fails to reach the ovule, indicating that the lower pollen tube growth in YS than in SY is one of the reasons of UCI in YS. The notable transcriptome change that is relevant to this microscopic observation is the expression of 4CLs and CA3OM. 4CLs controlling the essential steps of the biosynthesis of secondary metabolites leading to lignin (Hamberger and Hahlbrock, 2004). Their expression, together with CA3OM, indicates that lignin or other downstream metabolites in YS24 must be decreased, whereas in SY it was almost normal. This is relatable to the known fact that a massive cell wall deposition is required for fast elongation of pollen tubes, and that we observed faint pollen tube signals in YS24 (**Figure 2**). Earlier studies explained that the deposition of lignin was affected by changes in the expression of 4CLs and CA3OM (Ma and Xu, 2008; Cao et al., 2020). Thus, these genes would be a suitable candidate for increasing lignin deposition in elongating pollen tubes.

The expression changes in CESAs suggests two things, i.e., first, cellulose synthesis in YS is affected similar to the lignin biosynthesis and, second, that YS24 may have lower cellulose deposition than SY24 (Nawaz et al., 2019). The second observation could not be generalized, since two other CESAs showed higher expression in YS24 than in CK1, and their expression pattern was similar in both pollination combinations. The observation that there were three CESA transcripts could be due to the fact that there are multiple CESA members that differ in function and are expressed differently (Nawaz et al., 2017a, 2019). Furthermore, in Arabidopsis, the silencing of *CESA1* and *CESA*3 caused no pollen tube production (Persson et al., 2007). Only a further experiment might reveal if *Dimocarpus_longan_newGene_10844* or the other two also have a possible role in UCI in the YS pollination combination. Xyloglucan is a major hemicellulosic polysaccharide of a primary cell wall, which interacts with cellulose and controls cell wall expansion. Our observation that the expression of most of the XETs was higher in YS24 than in CK1 and all the SY time points is in line with earlier reports that XET activity loosens plant cell walls (Van Sandt et al., 2007; Nawaz et al., 2017b). It suggests, that in the YS pollination combination, the pollen tube could be under modification higher than that of the SY pollination combination. This proposition is further supported by the expression pattern observed for XIPs (*Dlo_001017.1.gene* and *Dlo_022672.1.gene*), as XIPs are known for degradation of xylans and the fact that in rice they are downregulated during pollen tube growth (Dai et al., 2007). In addition, it is known that the application (or increased expression) of PMEs causes a dramatic decrease in pollen tube germination/growth (Bosch et al., 2005). One PME (also known as PE, *Dlo_023530.1.gene*) showed higher expression in YS24 (8.58) than in SY24 (3.06). The expression pattern of one PME (*Dlo_023530.1.gene*) is consistent with the known role of PMEs in *Nicotiana tabacum* (Bosch et al., 2005). In accordance with this, the differential expression of PMEI (*Dlo_024765.1.gene*) in YS and SY compared to CK1 and CK2, respectively, indicates that in YS the PME level might be higher because of the increased expression of PME/PE as a result of the reduced expression of PMEI and vice versa in SY. We say this because PMEI inhibits PME from flower, siliques, and pollens (Wolf et al., 2003). Taken together, our transcriptome results show that genes associated

with cell wall biosynthesis, modification, and lignin deposition express differentially between both YS and SY pollination combinations, and are one of the reasons for UCI in YS. These results lay down the foundations of specific future investigations targeting cell wall modification during YS pollen tube growth compared to that of SY.

Differential Expression of Pollen Tube Growth-Related Genes Was Observed Between Shixia♀ × Yiduo♂ and Yiduo♀ × Shixia♂

Pollen tubes are fast-growing cells which require multiple factors. *In vitro* pollen tube growth studies have highlighted that Ca^{2+}, K^{+}, and pH dynamics are prime factors that accompany tip growth (Hepler et al., 2006; Scheible and McCubbin, 2019). Thus, ion transporters are prime targets to understand if there are differences in pollen tube growth rate, as observed in our study (**Figures 2**, **3**). The observation that a CHX15 was upregulated in the SY series compared to CK2 and downregulated in the YS series compared to CK1 is relevant to recent findings that two CHXs (CHX21 and CHX23), when mutated, result in normal pollen germination but are defective in pollen tube guidance to the ovule (Lu et al., 2011). CHX15, CHX21, and CHX23 belong to same subfamily of cation exchangers [see UniProtKB - Q9SIT5[4]; (Chanroj et al., 2011)]; therefore, a similar role can be expected. This is an interesting observation and a good gene of interest for a detailed exploration of its role in longan UCI. Studies have elaborated that K^{+} is required to regulate turgor pressure in pollen tubes to burst during fertilization (Holdaway-Clarke and Hepler, 2003), and that K^{+} influx promotes pollen tube growth (Mouline et al., 2002). At a relatively higher Ca^{2+} concentration, the mutation of CPKs did not inhibit pollen tube growth in Arabidopsis, indicating that CPKs (CPK11 and CPK24) control pollen tube growth. Our observations that the expression of two CPK17s was higher in YS than in CK1 and lower in SY than in CK2 can be relevant to this known fact. Particularly, the higher CPK24 expression in YS24 than in SY24 is consistent with this report (Zhao et al., 2013). Ca^{2+} channels are core components in pollen tube tips that regulate Ca^{2+} gradients. Mutations in CNGC18 have been studied in Arabidopsis pollen tube growth and were found critically responsible for pollen tube guidance. The similar expression trend of the CNGC18 (Gao et al., 2016) indicates its lesser role in UCI in the YS longan pollination combination. Another protein whose localization is influenced by Ca^{2+} is syntaxin (Silva et al., 2010). In our data, the higher expression of SYP124 in YS than in SY is consistent with an earlier report as well as the expression pattern of CPKs (Silva et al., 2010). The expression of 1-phosphatidylinositol-4-phosphate 5-kinase is also consistent with the known function that homozygous mutants lacking *PIP5K4* transcript significantly affect pollen tube growth and polarity (Sousa et al., 2008). Finally, pollen tube growth has also been related with Arabidopsis inositol transporter 4 (INT4) (Schneider et al., 2006). The expression of INT13 is consistent with the expression in Arabidopsis that INT4 was detected in growing pollen tubes and plays a role in high rate of pollen rate growth (Schneider et al., 2006). Taken together, our observations from transcriptome sequencing and comparison indicate that many genes associated with pollen tube growth were differentially expressed between SY and YS and cannot be ignored particularly in relation to the microscopic observations made in our report.

Starch and Sucrose Biosynthesis Pathway Is Differentially Regulated in the Yiduo♀ × Shixia♂ and Shixia♀ × Yiduo♂ Pollination Combinations

Fast-growing pollen tubes require different resources such as sugars, cell wall polymers, and energy. It has been established that sucrose content is positively correlated with pollen germination as well as pollen tube growth in *N. tabacum* (Parrotta et al., 2018) and *Cunnighamial lanceolata* L (Fragallah et al., 2019). Thus, regulation of the starch and glucose biosynthesis pathway can highlight potential genes that play a role in UCI in the YS longan pollination combination. The downregulation of SuSy and SPS in YS compared to SY is consistent with our observations on pollen tube growth rate (**Figures 2**, **3**) and changes in the expression of genes related to pollen tube and cell wall (see above sections). Both the SuSy and SPS genes are required for sucrose metabolism (Winter and Huber, 2000), and their expression changes indicate that this process might be disturbed in the YS pollination combination compared to SY. The expression of HKs and FRKs indicate that in YS, the conversion of D-fructose to D-fructose-6P is reduced compared to the SY pollination combination. This statement is based on the known function of both enzymes (Jiang et al., 2003; Kawai et al., 2005). However, transcripts for a gene present upstream this step, i.e., BFRUCTF, showed decreased expression in both pollination combinations, indicating depletion of sucrose (Schwebel-Dugué et al., 1994), which shows that sucrose breakdown is similar in both pollination combinations. The breakdown of sucrose also results in D-glucose (Schwebel-Dugué et al., 1994). In this regard, we noted the differential expression of multiple genes that take part in D-glucose biosynthesis in the starch and sucrose biosynthesis pathway. The breakdown of cellulose to cellodextin by endoglucanase (EG, *Dlo_022739.1.gene* and *Dlo_028821.1.gene*) (Opassiri et al., 2006) and then to cellobiose by EG and beta-glucosidases (BGLs) (Yoshida and Komae, 2006) leads to D-glucose biosynthesis. The variable expressions of EGs and BGLs enable us to think that there are changes happening in cellulose and its conversion to D-glucose. Apart from sucrose and D-glucose metabolism-related transcriptome changes, the fact that multiple genes controlling starch metabolism, i.e., 1,4-aplha-glucan branching enzyme, AAs, and BAs were differentially regulated between YS and SY, indicating changes in starch concentration in the studied tissues. Most probably, we can understand that starch/glycogen biosynthesis in YS must be higher than in SY, and that its conversion to maltose and/or destrin is

[4]https://www.uniprot.org/uniprot/Q9SIT5

also higher in YS than in SY. However, these expression changes cannot be considered a direct cause of UCI in YS in comparison to SY.

Jasmonic Acid Metabolism and Signaling Are Regulated Differently in the Yiduo♀ × Shixia♂ and Shixia♀ × Yiduo♂ Pollination Combinations

Since MeJA is known to affect pollen germination as well as pollen tube growth (Muradoğlu et al., 2010), we specifically explored the JA metabolism and signaling pathways. Muradoğlu, Yıldız and Balta (Muradoğlu et al., 2010) reported that MeJA led to a decrease in pollen germination as well as pollen tube growth in apricots. Another study reported that the application of 2.5 mM MeJA resulted in no pollen germination in *Pinus nigra* (Çetinbaş-Genç and Vardar, 2020). Our observations that both JA metabolism and signaling were differentially regulated in both the YS and SY pollination combinations are in accordance with these reports. Most importantly, in JA metabolism, the conversion of galactolipids into α-linoleic acid is controlled by PLA1 and DAD1 (Ishiguro et al., 2001). The expression changes in PLA1 indicate that the conversion of galactolipids into α-linoleic acid is higher in SY and relatively lower in YS. This was further confirmed by the expression trend of PLA1, as noted in the RT-qPCR analysis (**Supplementary Figure 3**). Next step that leads the biosynthesis of JA is the LOX pathway, i.e., AOS branch (Feussner and Wasternack, 2002). The expression patterns of genes that control important steps of the LOX pathway, i.e., LOXs, OPR11, and JMT in YS (particularly at YS24) compared to SY are consistent with observed changes in endogenous JA and JA-Ile levels based on the known functions of the respective genes (also see **Supplementary Figure 3**) (Ishiguro et al., 2001; Seo et al., 2001; Feussner and Wasternack, 2002; Wasternack and Hause, 2013). JA-Ile is formed by JAR1, which is a member of the GH3 gene family (Staswick and Tiryaki, 2004). This is a major step in JA perception, since JA-Ile is the most biologically active JA compound (Fonseca et al., 2009). Its higher expression in YS is consistent with the higher JA-Ile concentration (**Figure 7** and **Supplementary Figure 3**). JA-Ile signal is perceived by COI, which triggers the repression of JAZ. The transcriptome comparisons showed that JA-Ile triggered the repression of JAZs in both YS and SY, although we did not note the differential regulation of COI. JAZ repression releases downstream transcription factors (TFs), e.g., MYC2s. The expression pattern of MYC2s indicates that in the YS pollination combination, the JA-responsive activation of MYC2 is relatively lower than that in SY (in particular 12 and 24 HAP) (**Supplementary Figure 3**). There are multiple genes present downstream of MYC2s that regulate JA-triggered responses in plant tissues (Ruan et al., 2019). These results are in accordance with the fact that the downregulation of these genes has been linked with the inhibition of endogenous JA synthesis in Arabidopsis during pollen germination (Ju et al., 2016). Taken together, the changes in the expressions of JA biosynthesis- and signaling-related genes indicate an important role of JA in UCI. However, further studies involving exogenous application of JA/MeJA or agents that can regulate its *in planta* biosynthesis will elaborate its role. Also, characterization of genes through gene silencing or other techniques would be useful to explain clearly how the expression of the individual genes affect UCI.

CONCLUSION

The reciprocal crossing in longan cultivars Yiduo and Shixia is successful only in one way, i.e., SY, while in the case of YS, it is incompatible. Our results elaborated the differential expression of genes associated with plant-hormone signal transduction, phenylpropanoid biosynthesis, protein processing in the endoplasmic reticulum, and starch and sucrose biosynthesis pathways. The detailed analysis of plant-hormone signaling pathway indicated that JA metabolism and signaling have important roles in UCI. The endogenous JA contents in the YS pollination combination were higher than in SY. In the case of YS, pollen tube growth was slow, and the transcriptome comparison data could be related to it by discussing pollen tube, cell wall, and starch and sucrose metabolism-related transcripts. The candidate genes identified in this study represent key resources for functional studies related to UCI in longan.

DATA AVAILABILITY STATEMENT

The raw RNA-seq data has been submitted to NCBI SRA and released under the project number PRJNA741615.

AUTHOR CONTRIBUTIONS

JW collected the data, and analyzed and drafted the article. DG did the preparation and treatment of experimental materials. JW, DG, DH, and SH revised the manuscript. JC did the microscopic observation. DG and JL guided the experiment, provided funding, and revised the article. All authors have read and approved the final version of this manuscript.

FUNDING

This study was funded by Key Area Research and Development Program of Guangdong Province (No. 2020B020220006), Guangdong Academy of Agricultural Sciences Foundation (No. 201916), and Guangdong Provincial Crops Germplasm Nursery Construction and Resources Collection, Preservation, Identification, and Evaluation Foundation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.821147/full#supplementary-material

REFERENCES

Abdalla, M. M. (1977). Intraspecific unilateral incompatibility in *Vicia faba* L. *Theor. Appl. Genet.* 50, 227–233. doi: 10.1007/BF00273756

Adhikari, P. B., Liu, X., and Kasahara, R. D. (2020). Mechanics of pollen tube elongation: a perspective. *Front. Plant Sci.* 11:589712. doi: 10.3389/fpls.2020.589712

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389

Apweiler, R., Bairoch, A., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., et al. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* 32, D115–D119. doi: 10.1093/nar/gkh131.10.1093/nar/gkw1099

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Baecker, P. A., Greenberg, E., and Preiss, J. (1986). Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* 1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-alpha-D-(1, 4-alpha-D-glucano)-transferase as deduced from the nucleotide sequence of the glg B gene. *J. Biol. Chem.* 261, 8738–8743.

Becker, J. D., and Feijó, J. A. (2007). How many genes are needed to make a pollen tube? Lessons from transcriptomics. *Ann. Bot.* 100, 1117–1123. doi: 10.1093/aob/mcm208

Borassi, C., Sede, A. R., Mecchia, M. A., Mangano, S., Marzol, E., Denita-Juarez, S. P., et al. (2021). Proline-rich extensin-like receptor kinases PERK5 and PERK12 are involved in pollen tube growth. *FEBS Lett.* 595, 2593–2607. doi: 10.1002/1873-3468.14185

Bosch, M., Cheung, A. Y., and Hepler, P. K. (2005). Pectin methylesterase, a regulator of pollen tube growth. *Plant Physiol.* 138, 1334–1346. doi: 10.1104/pp.105.059865

Cao, S., Huang, C., Luo, L., Zheng, S., Zhong, Y., Sun, J., et al. (2020). Cell-specific suppression of 4-coumarate-CoA ligase gene reveals differential effect of lignin on cell physiological function in *Populus*. *Front. Plant Sci.* 11:589729. doi: 10.3389/fpls.2020.589729

Caser, M. (2017). "Pollen grains and tubes," in *Reference Module in Life Sciences*, ed. M. Caser (Amsterdam: Elsevier).

Çetinbaş-Genç, A., and Vardar, F. (2020). Effect of methyl jasmonate on in-vitro pollen germination and tube elongation of *Pinus nigra*. *Protoplasma* 257, 1655–1665. doi: 10.1007/s00709-020-01539-4

Chanroj, S., Lu, Y., Padmanaban, S., Nanatani, K., Uozumi, N., Rao, R., et al. (2011). Plant-specific cation/H+ exchanger 17 and its homologs are endomembrane K+ transporters with roles in protein sorting. *J. Biol. Chem.* 286, 33931–33941. doi: 10.1074/jbc.M111.252650

Chaturvedi, P., Ghatak, A., and Weckwerth, W. (2016). Pollen proteomics: from stress physiology to developmental priming. *Plant Reprod.* 29, 119–132. doi: 10.1007/s00497-016-0283-9

Choo, W. K. (2000). *Longan Production in Asia*. Rome: FAO.

Dai, S., Chen, T., Chong, K., Xue, Y., Liu, S., and Wang, T. (2007). Proteomics identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated *Oryza sativa* pollen. *Mol. Cell Proteomics* 6, 207–230. doi: 10.1074/mcp.M600146-MCP200

De Nettancourt, D. (2001). *Incompatibility and Incongruity in Wild and Cultivated Plants*. Cham: Springer Science & Business Media.

Deng, Y., Li, J., Wu, S., Zhu, Y., Chen, Y., and He, F. (2006). Integrated nr database in protein annotation system and its localization. *Comput. Eng.* 32, 71–72.

Deng, Y., Sun, X., Gu, C., Jia, X., Liang, L., and Su, J. (2017). Identification of pre-fertilization reproductive barriers and the underlying cytological mechanism in crosses among three petal-types of Jasminum sambac and their relevance to phylogenetic relationships. *PLoS One* 12:e0176026. doi: 10.1371/journal.pone.0176026

Feussner, I., and Wasternack, C. (2002). The lipoxygenase pathway. *Annu. Rev. Plant Biol.* 53, 275–297.

Fonseca, S., Chini, A., Hamberg, M., Adie, B., Porzel, A., Kramell, R., et al. (2009). (+)-7-iso-Jasmonoyl-L-isoleucine is the endogenous bioactive jasmonate. *Nat. Chem. Biol.* 5, 344–350. doi: 10.1038/nchembio.161

Fragallah, S. A. D. A., Lin, S., Li, N., Ligate, E. J., and Chen, Y. (2019). Effects of sucrose, boric acid, pH, and incubation time on in vitro germination of pollen and tube growth of chinese fir (*Cunnighamial lanceolata* L.). *Forests* 10:102.

Gao, Q. F., Gu, L. L., Wang, H. Q., Fei, C. F., Fang, X., Hussain, J., et al. (2016). Cyclic nucleotide-gated channel 18 is an essential Ca2+ channel in pollen tube tips for pollen tube guidance to ovules in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3096–3101. doi: 10.1073/pnas.1524629113

Ghosh, P., Meyer, C., Remy, E., Peterson, D., and Preiss, J. (1992). Cloning, expression, and nucleotide sequence of glgC gene from an allosteric mutant of *Escherichia coli* B. *Arch. Biochem. Biophys.* 296, 122–128. doi: 10.1016/0003-9861(92)90553-9

Guan, Y., Guo, J., Li, H., and Yang, Z. (2013). Signaling in pollen tube growth: crosstalk, feedback, and missing links. *Mol. Plant* 6, 1053–1064. doi: 10.1093/mp/sst070

Hamberger, B., and Hahlbrock, K. (2004). The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 2209–2214. doi: 10.1073/pnas.0307307101

Hasachoo, N., and Kalaya, P. (2013). *Competitiveness of Local Agriculture: The Case of Longan Fruit Trade Between China and the North of Thailand*. Bangkok: The Research Institute on Contemporary Southeast Asia (IRASEC).

He, Y., Song, Q., Chen, S., Wu, Y., Zheng, G., Feng, J., et al. (2020). Transcriptome analysis of self-and cross-pollinated pistils revealing candidate unigenes of self-incompatibility in *Camellia oleifera*. *J. Hortic. Sci. Biotech.* 95, 19–31.

Hepler, P. K., Lovy-Wheeler, A., McKenna, S. T., and Kunkel, J. G. (2006). Ions and pollen tube growth. *Pollen Tube* 3, 47–69.

Holdaway-Clarke, T. L., and Hepler, P. K. (2003). Control of pollen tube growth: role of ion gradients and fluxes. *New Phytol.* 159, 539–563. doi: 10.1046/j.1469-8137.2003.00847.x

Holmes, G. D., James, E. A., and Hoffmann, A. A. (2008). Limitations to reproductive output and genetic rescue in populations of the rare shrub *Grevillea repens* (Proteaceae). *Ann. Bot.* 102, 1031–1041. doi: 10.1093/aob/mcn195

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* 44, D286–D293. doi: 10.1093/nar/gkv1248

Iaria, D., Chiappetta, A., and Muzzalupo, I. (2016). De novo transcriptome sequencing of *Olea europaea* L. to identify genes involved in the development of the pollen tube. *ScientificWorldJournal* 2016:4305252. doi: 10.1155/2016/4305252

Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I., and Okada, K. (2001). The defective in anther dehiscience gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *Plant Cell* 13, 2191–2209. doi: 10.1105/tpc.010192

Jiang, H., Dian, W., Liu, F., and Wu, P. (2003). Isolation and characterization of two fructokinase cDNA clones from rice. *Phytochemistry* 62, 47–52. doi: 10.1016/s0031-9422(02)00428-4

Ju, Y., Guo, L., Cai, Q., Ma, F., Zhu, Q. Y., Zhang, Q., et al. (2016). Arabidopsis JINGUBANG is a negative regulator of pollen germination that prevents pollination in moist environments. *Plant Cell* 28, 2131–2146. doi: 10.1105/tpc.16.00401

Jue, D., Sang, X., Liu, L., Shu, B., Wang, Y., Liu, C., et al. (2019). Comprehensive analysis of the longan transcriptome reveals distinct regulatory programs during the floral transition. *BMC Genomics* 20:126. doi: 10.1186/s12864-019-5461-3

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32, D277–D280. doi: 10.1093/nar/gkh063

Kawai, S., Mukai, T., Mori, S., Mikami, B., and Murata, K. (2005). Hypothesis: structures, evolution, and ancestor of glucose kinases in the hexokinase family. *J. Biosci. Bioeng.* 99, 320–330. doi: 10.1263/jbb.99.320

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Koonin, E. V., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Krylov, D. M., Makarova, K. S., et al. (2004). A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* 5:R7. doi: 10.1186/gb-2004-5-2-r7

Krichevsky, A., Kozlovsky, S. V., Tian, G.-W., Chen, M. H., Zaltsman, A., and Citovsky, V. (2007). How pollen tubes grow. *Dev. Biol.* 303, 405–420.

Lee, S. I., Muthusamy, M., Nawaz, M. A., Hong, J. K., Lim, M.-H., Kim, J. A., et al. (2019). Genome-wide analysis of spatiotemporal gene expression patterns during floral organ development in *Brassica rapa*. *Mol. Genet. Genom.* 294, 1403–1420. doi: 10.1007/s00438-019-01585-5

Lin, Y., Min, J., Lai, R., Wu, Z., Chen, Y., Yu, L., et al. (2017). Genome-wide sequencing of longan (*Dimocarpus longan* Lour.) provides insights into molecular basis of its polyphenol-rich characteristics. *Gigascience* 6, 1–14. doi: 10.1093/gigascience/gix023

López-Cristoffanini, C., Serrat, X., Jáuregui, O., Nogués, S., and López-Carbonell, M. (2019). Phytohormone profiling method for rice: effects of GA20ox mutation on the gibberellin content of japonica rice varieties. *Front. Plant Sci.* 10:733. doi: 10.3389/fpls.2019.00733

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lu, Y., Chanroj, S., Zulkifli, L., Johnson, M. A., Uozumi, N., Cheung, A., et al. (2011). Pollen tubes lacking a pair of K+ transporters fail to target ovules in *Arabidopsis*. *Plant Cell* 23, 81–93. doi: 10.1105/tpc.110.080499

Luo, T., Shuai, L., Lai, T., Liao, L., Li, J., Duan, Z., et al. (2021). Up-regulated glycolysis, TCA, fermentation and energy metabolism promoted the sugar receding in 'Shixia'longan (*Dimocarpus longan* Lour.) pulp. *Sci. Hortic.* 281:109998.

Ma, Q. H., and Xu, Y. (2008). Characterization of a caffeic acid 3-O-methyltransferase from wheat and its function in lignin biosynthesis. *Biochimie* 90, 515–524. doi: 10.1016/j.biochi.2007.09.016

Majid, R. (1964). Overcoming interspecific incompatibility in *Lycopersicon*. *Curr. Sci. India* 33, 154–156.

Maruyama, D., Endo, T., and Nishikawa, S. (2010). BiP-mediated polar nuclei fusion is essential for the regulation of endosperm nuclei proliferation in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1684–1689. doi: 10.1073/pnas.0905795107

Maruyama, D., Sugiyama, T., Endo, T., and Nishikawa, S. (2014). Multiple BiP genes of *Arabidopsis thaliana* are required for male gametogenesis and pollen competitiveness. *Plant Cell Physiol.* 55, 801–810. doi: 10.1093/pcp/pcu018

Mascarenhas, J. P. (1993). Molecular mechanisms of pollen tube growth and differentiation. *Plant Cell* 5, 1303–1314. doi: 10.1105/tpc.5.10.1303

McConchie, C., Vithanage, V., and Batten, D. (1994). Intergeneric hybridisation between litchi (*Litchi chinensis* Sonn.) and longan (*Dimocarpus longan* Lour.). *Ann. Bot.* 74, 111–118.

Mollet, J.-C., Leroux, C., Dardelle, F., and Lehner, A. (2013). Cell wall composition, biosynthesis and remodeling during pollen tube growth. *Plants* 2, 107–147. doi: 10.3390/plants2010107

Mouline, K., Véry, A.-A., Gaymard, F., Boucherez, J., Pilot, G., Devic, M., et al. (2002). Pollen tube development and competitive ability are impaired by disruption of a Shaker K(+) channel in *Arabidopsis*. *Genes Dev.* 16, 339–350. doi: 10.1101/gad.213902

Muradoğlu, F., Yıldız, K., and Balta, F. (2010). Methyl jasmonate influences of pollen germination and pollen tube growth of apricot (*Prunus armeniaca* L.). *Yüzüncü Yıl Univ. J. Agric. Sci.* 20, 183–188.

Nawaz, M. A., Lin, X., Chan, T. F., Imtiaz, M., Rehman, H. M., Ali, M. A., et al. (2019). Characterization of cellulose synthase A (CESA) gene family in eudicots. *Biochem. Genet.* 57, 248–272. doi: 10.1007/s10528-018-9888-z

Nawaz, M. A., Rehman, H. M., Baloch, F. S., Ijaz, B., Ali, M. A., Khan, I. A., et al. (2017a). Genome and transcriptome-wide analyses of cellulose synthase gene superfamily in soybean. *J. Plant Physiol.* 215, 163–175. doi: 10.1016/j.jplph.2017.04.009

Nawaz, M. A., Rehman, H. M., Imtiaz, M., Baloch, F. S., Lee, J. D., Yang, S. H., et al. (2017b). Systems identification and characterization of cell wall reassembly and degradation related genes in *Glycine max* (L.) merill, a bioenergy legume. *Sci. Rep.* 7:10862. doi: 10.1038/s41598-017-11495-4

O'Kelley, J. C. (1955). External carbohydrates in growth and respiration of pollen tubes in vitro. *Am. J. Bot.* 42, 322–327.

Onus, A. N., and Pickersgill, B. (2004). Unilateral incompatibility in *Capsicum* (Solanaceae): occurrence and taxonomic distribution. *Ann. Bot.* 94, 289–295. doi: 10.1093/aob/mch139

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Cairns, J. R. K. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 beta-glucosidase. *BMC Plant Biol.* 6:33. doi: 10.1186/1471-2229-6-33

Parrotta, L., Faleri, C., Del Duca, S., and Cai, G. (2018). Depletion of sucrose induces changes in the tip growth mechanism of tobacco pollen tubes. *Ann. Bot.* 122, 23–43. doi: 10.1093/aob/mcy043

Persson, S., Paredez, A., Carroll, A., Palsdottir, H., Doblin, M., Poindexter, P., et al. (2007). Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 104, 15566–15571. doi: 10.1073/pnas.0706592104

Pham, H. D. (2012). *Pollination Biology of Jujubes and Longans and the Importance of Insects in the Pollination of Crops in Vietnam*. Guelph: University of Guelph.

Qiu, D. (2012). "Longan production and research in China," in *Proceedings of IV International Symposium on Lychee, Longan and Other Sapindaceae Fruits*, Hanoi, 39–46.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nat. Biotechnol.* 29, 24–26.

Ruan, J., Zhou, Y., Zhou, M., Yan, J., Khurshid, M., Weng, W., et al. (2019). Jasmonic acid signaling pathway in plants. *Int. J. Mol. Sci.* 20:2479. doi: 10.3390/ijms20102479

Scali, M., Moscatelli, A., Bini, L., Onelli, E., Vignani, R., and Wang, W. (2021). Protein analysis of pollen tubes after the treatments of membrane trafficking inhibitors gains insights on molecular mechanism underlying pollen tube polar growth. *Protein J.* 40, 205–222. doi: 10.1007/s10930-021-09972-x

Schaeffer, G. W., Haskins, F. A., and Gorz, H. J. (1960). Genetic control of coumarin biosynthesis and beta-glucosidase activity in *Melilotus alba*. *Biochem. Biophys. Res. Commun.* 3, 268–271. doi: 10.1016/0006-291x(60)90237-0

Scheible, N., and McCubbin, A. (2019). Signaling in pollen tube growth: beyond the tip of the polarity iceberg. *Plants* 8:156. doi: 10.3390/plants8060156

Schneider, S., Schneidereit, A., Konrad, K. R., Hajirezaei, M. R., Gramann, M., Hedrich, R., et al. (2006). *Arabidopsis* INOSITOL TRANSPORTER4 mediates high-affinity H+ symport of myoinositol across the plasma membrane. *Plant Physiol.* 141, 565–577. doi: 10.1104/pp.106.077123

Schwebel-Dugué, N., El Mtili, N., Krivitzky, M., Jean-Jacques, I., Williams, J. H., Thomas, M., et al. (1994). *Arabidopsis* gene and cDNA encoding cell-wall invertase. *Plant Physiol.* 104, 809–810. doi: 10.1104/pp.104.2.809

Seo, H. S., Song, J. T., Cheong, J. J., Lee, Y. H., Lee, Y. W., Hwang, I., et al. (2001). Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4788–4793. doi: 10.1073/pnas.081557298

Sheoran, I. S., Ross, A. R., Olson, D. J., and Sawhney, V. K. (2007). Proteomic analysis of tomato (*Lycopersicon esculentum*) pollen. *J. Exp. Bot.* 58, 3525–3535. doi: 10.1093/jxb/erm199

Silva, P. Â, Ul-Rehman, R., Rato, C., Di Sansebastiano, G. P., and Malhó, R. (2010). Asymmetric localization of *Arabidopsis* SYP124 syntaxin at the pollen tube apical and sub-apical zones is involved in tip growth. *BMC Plant Biol.* 10:179. doi: 10.1186/1471-2229-10-179

Sousa, E., Kost, B., and Malhó, R. (2008). Arabidopsis phosphatidylinositol-4-monophosphate 5-kinase 4 regulates pollen tube growth and polarity by modulating membrane recycling. *Plant Cell* 20, 3050–3064. doi: 10.1105/tpc.108.058826

Stadler, R., Truernit, E., Gahrtz, M., and Sauer, N. (1999). The AtSUC1 sucrose carrier may represent the osmotic driving force for anther dehiscence and pollen tube growth in *Arabidopsis*. *Plant J.* 19, 269–278. doi: 10.1046/j.1365-313x.1999.00527.x

Staswick, P. E., and Tiryaki, I. (2004). The oxylipin signal jasmonic acid is activated by an enzyme that conjugates it to isoleucine in *Arabidopsis*. *Plant Cell* 16, 2117–2127. doi: 10.1105/tpc.104.023549

Stone, L., Seaton, K., Kuo, J., and McComb, J. (2004). Fast pollen tube growth in *Conospermum* species. *Ann. Bot.* 93, 369–378. doi: 10.1093/aob/mch050

Subhadrabandhu, S., and Yapwattanaphun, C. (2001). "Lychee and longan production in Thailand," in *Proceedings of I International Symposium on Litchi and Longan*, Hanoi, 49–57.

Takada, Y., Murase, K., Shimosato-Asano, H., Sato, T., Nakanishi, H., Suwabe, K., et al. (2017). Duplicated pollen–pistil recognition loci control intraspecific unilateral incompatibility in *Brassica rapa*. *Nat. Plants* 26:17096. doi: 10.1038/nplants.2017.96

Tatusov, R. L., Galperin, M. Y., Natale, D. A., and Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36. doi: 10.1093/nar/28.1.33

Tovar-Méndez, A., Kumar, A., Kondo, K., Ashford, A., Baek, Y. S., Welch, L., et al. (2014). Restoring pistil-side self-incompatibility factors recapitulates an interspecific reproductive barrier between tomato species. *Plant J.* 77, 727–736. doi: 10.1111/tpj.12424

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35, W71–W74. doi: 10.1093/nar/gkm306

Van Creij, M., Kerckhoffs, D., and Van Tuyl, J. (1997). "Application of three pollination techniques and of hormone treatments for overcoming interspecific crossing barriers in *Tulipa*," in *Proceedings of VII International Symposium on Flowerbulbs 430*, ed. A. H. Halevy (Leiden: International Society for Horticultural Science), 547–558.

Van Creij, M., Kerckhoffs, D., and Van Tuyl, J. (2000). "Application of four pollination techniques and of hormone treatment for bypassing interspecific crossing barriers in *Lilium* L," in *Proceedings of XIX International Symposium on Improvement of Ornamental Plants 508*, ed. A. Cadic (Leiden: International Society for Horticultural Science), </UB>267–276.

Van Sandt, V. S., Suslov, D., Verbelen, J. P., and Vissenberg, K. (2007). Xyloglucan endotransglucosylase activity loosens a plant cell wall. *Ann. Bot.* 100, 1467–1473. doi: 10.1093/aob/mcm248

Wang, M., Chen, Z., Zhang, H., Chen, H., and Gao, X. (2018). Transcriptome analysis provides insight into the molecular mechanisms underlying gametophyte factor 2-mediated cross-incompatibility in maize. *Int. J. Mol. Sci.* 19:1757. doi: 10.3390/ijms19061757

Wasternack, C., and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Ann. Bot.* 111, 1021–1058. doi: 10.1093/aob/mct067

Winter, H., and Huber, S. C. (2000). Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *CRC Crit. Rev. Plant Sci.* 19, 31–67.

Wolf, S., Grsic-Rausch, S., Rausch, T., and Greiner, S. (2003). Identification of pollen-expressed pectin methylesterase inhibitors in *Arabidopsis*. *FEBS Lett.* 555, 551–555. doi: 10.1016/s0014-5793(03)01344-9

Xiao, H. M., Cai, W. J., Ye, T. T., Ding, J., and Feng, Y. Q. (2018). Spatio-temporal profiling of abscisic acid, indoleacetic acid and jasmonic acid in single rice seed during seed germination. *Anal. Chim. Acta* 1031, 119–127. doi: 10.1016/j.aca.2018.05.055

Xie, D. R., Ma, X. S., Rahman, M. Z., Yang, M. C., Huang, X. M., Li, J. G., et al. (2019). Thermo-sensitive sterility and self-sterility underlie the partial seed abortion phenotype of *Litchi chinensis*. *Sci. Hortic.* 247, 156–164.

Yang, J., Duan, G., Li, C., Liu, L., Han, G., Zhang, Y., et al. (2019). The crosstalks between jasmonic acid and other plant hormone signaling highlight the involvement of jasmonic acid as a core component in plant response to biotic and abiotic stresses. *Front. Plant Sci.* 10:1349. doi: 10.3389/fpls.2019.01349

Yoshida, K., and Komae, K. (2006). A rice family 9 glycoside hydrolase isozyme with broad substrate specificity for hemicelluloses in type II cell walls. *Plant Cell Physiol.* 47, 1541–1554. doi: 10.1093/pcp/pcl020

Zhang, C. C., Wang, L. Y., Wei, K., Wu, L. Y., Li, H. L., Zhang, F., et al. (2016). Transcriptome analysis reveals self-incompatibility in the tea plant (*Camellia sinensis*) might be under gametophytic control. *BMC Genomics* 17:359. doi: 10.1186/s12864-016-2703-5

Zhang, S., Ding, F., He, X., Luo, C., Huang, G., and Hu, Y. (2015). Characterization of the 'Xiangshui' lemon transcriptome by de novo assembly to discover genes associated with self-incompatibility. *Mol. Genet. Genomics* 290, 365–375. doi: 10.1007/s00438-014-0920-7

Zhang, X., and Liu, C. J. (2015). Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids. *Mol. Plant* 8, 17–27. doi: 10.1016/j.molp.2014.11.001

Zhao, L. N., Shen, L. K., Zhang, W. Z., Zhang, W., Wang, Y., and Wu, W. H. (2013). Ca2+-dependent protein kinase11 and 24 modulate the activity of the inward rectifying K+ channels in *Arabidopsis* pollen tubes. *Plant Cell* 25, 649–661. doi: 10.1105/tpc.112.103184

Zou, T., Li, S., Liu, M., Wang, T., Xiao, Q., Chen, D., et al. (2017). An atypical strictosidine synthase, OsSTRL2, plays key roles in anther development and pollen wall formation in rice. *Sci. Rep.* 7:6863. doi: 10.1038/s41598-017-07064-4

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fpls.2022.834431

Downregulation of miR156-Targeted *PvSPL6* in Switchgrass Delays Flowering and Increases Biomass Yield

Jinjun Cai[1,2], Wenwen Liu[3], Weiqian Li[2], Lijuan Zhao[3,4], Gang Chen[2], Yangyang Bai[2], Dongmei Ma[4], Chunxiang Fu[3,5], Yamei Wang[3] and Xinchang Zhang[1]**

[1]College of Natural Resources and Environment, Northwest Agriculture and Forestry University, Yangling, China, [2]Institute of Agricultural Resources and Environment, Ningxia Academy of Agriculture and Forestry Sciences, Yinchuan, China, [3]Shandong Provincial Key Laboratory of Energy Genetics and CAS Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China, [4]Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwest China, Ningxia University, Yinchuan, China, [5]CAS Key Laboratory of Tibetan Medicine Research, Northwest Institute of Plateau Biology, Xining, China

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Shuizhang Fei,
Iowa State University, United States
Wanjun Zhang,
China Agricultural University, China
Ziliang Luo,
University of Florida, United States

***Correspondence:**
Xinchang Zhang
zhangxc@ms.iswc.ac.cn
Yamei Wang
wangym@qibebt.ac.cn

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

Received: 14 December 2021
Accepted: 28 January 2022
Published: 18 February 2022

Citation:
Cai J, Liu W, Li W, Zhao L, Chen G, Bai Y, Ma D, Fu C, Wang Y and Zhang X (2022) Downregulation of miR156-Targeted PvSPL6 in Switchgrass Delays Flowering and Increases Biomass Yield. Front. Plant Sci. 13:834431. doi: 10.3389/fpls.2022.834431

MiR156/*SQUAMOSA PROMOTER BINDING-LIKE*s (*SPL*s) module is the key regulatory hub of juvenile-to-adult phase transition as a critical flowering regulator. In this study, a miR156-targeted *PvSPL6* was identified and characterized in switchgrass (*Panicum virgatum* L.), a dual-purpose fodder and biofuel crop. Overexpression of *PvSPL6* in switchgrass promoted flowering and reduced internode length, internode number, and plant height, whereas downregulation of *PvSPL6* delayed flowering and increased internode length, internode number, and plant height. Protein subcellular localization analysis revealed that PvSPL6 localizes to both the plasma membrane and nucleus. We produced transgenic switchgrass plants that overexpressed a *PvSPL6-GFP* fusion gene, and callus were induced from inflorescences of selected PvSPL6-GFP$_{OE}$ transgenic lines. We found that the PvSPL6-GFP fusion protein accumulated mainly in the nucleus in callus and was present in both the plasma membrane and nucleus in regenerating callus. However, during subsequent development, the signal of the PvSPL6-GFP fusion protein was detected only in the nucleus in the roots and leaves of plantlets. In addition, PvSPL6 protein was rapidly transported from the nucleus to the plasma membrane after exogenous GA_3 application, and returned from the plasma membrane to nucleus after treated with the GA_3 inhibitor (paclobutrazol). Taken together, our results demonstrate that *PvSPL6* is not only an important target that can be used to develop improved cultivars of forage and biofuel crops that show delayed flowering and high biomass yields, but also has the potential to regulate plant regeneration in response to GA_3.

Keywords: *PvSPL6*, flowering, biomass yield, subcellular localization, gibberellin, switchgrass

INTRODUCTION

Flowering is the key physiological juncture at which the plant transitions from vegetative to reproductive growth, and flowering at the optimum time is important for plant growth and reproductive success. The control of flowering time is also critical for yield formation in cereal crops and biomass accumulation in biofuel crops (Wang and Ge, 2006). Switchgrass (*Panicum virgatum* L.) has been developed into a dedicated, herbaceous bioenergy crop (Fu et al., 2011), and biomass yield is a major target trait for the genetic improvement of switchgrass. The biomass yield of gramineous plants shows little increase after the transformation from vegetative to reproductive growth, as the nutrient supply flows primarily toward the inflorescence at this time (Casler, 2012). Therefore, the genetic manipulation of flowering time is a key approach for improving the architecture and biomass yield of switchgrass and other biofuel and forage crops (Wang and Ge, 2006; Johnson et al., 2017).

The molecular regulation of flowering time involves complex, synergistic regulation by exogenous and endogenous factors (Wang, 2014; Park et al., 2016; Campos-Rivero et al., 2017; Cho et al., 2017; Inoescu et al., 2017). In addition, mechanisms of flowering time regulation vary greatly among different plant species (Hill and Li, 2016). The regulatory mechanism of flowering time control in the model plant *Arabidopsis thaliana* has been studied extensively. It involves five major pathways: the photoperiodic, gibberellin, autonomous, vernalization, and age pathways. Among these five pathways, aging has received widely attention (Amasino and Michaels, 2010; Wang, 2014). The age pathway ensures that plants flower even under noninductive conditions (Poethig, 2009; Wang, 2014). MiR156 and its target *SQUAMOSA PROMOTER BINDING-LIKE* (*SPL*) genes constitute the key regulatory hub of the age pathway (Fu et al., 2012; Teotia and Tang, 2015). *SPL* genes encode a class of plant-specific transcription factors (TFs) and are conserved in monocots and eudicots (Klein et al., 1996; Yang et al., 2007). In *Arabidopsis*, *SPL* genes can be divided into three functionally distinct groups: (1) *SPL2*/*9*/*10*/*11*/*13*/*15* participate in developmental stage transition, *SPL9* and *SPL15* play a major role in these processes; (2) *SPL3*/*4*/*5* promote the floral meristem identity transition; and (3) *SPL6* does not have a major function in shoot morphogenesis, but may be important for certain physiological processes (Xu et al., 2016). *SPL* genes have been found to promote flowering mainly through three pathways: (1) SPL3/4/5 redundantly promote flowering through direct activation of *LEAFY* (*LFY*), *FRUITFULL* (*FUL*), and *AP1* (*LEAFY*; Yamaguchi et al., 2009); (2) SPL9 positively regulates the floral promoters *FUL*, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*), and *AGL-LIKE 42* (*AGL42*; Wang et al., 2009). SPL9 can also promote the transcription of downstream miR172 (Wu et al., 2009), thereby inhibiting the expression of *APETALA2-LIKE* genes (*TARGET OF EAT1*, *TARGET OF EAT2*, *SCHLAFMUTZE*, and *SCHNARCHZAPFEN*; Aukerman and Sakai, 2003; Jung et al., 2007). The *AP2-like* genes can inhibit the expression of the flowering induction gene *FLOWERING LOCUS T* (*FT*). *FT* is induced by the photoperiodic pathway and regulated by *FUL* and *SOC1* under long-day conditions (Litt and Irish, 2003; Mathieu et al., 2007); and (3) SPL2/10/11, which have close homology to SPL9, can affect the flowering process by regulating *FUL* gene expression (Wang et al., 2009). In contrast to the extensive studies in *Arabidopsis*, little information is available on the flowering-related roles of *SPL*s in the Gramineae. The miR156-targeted *PvSPL*s in switchgrass belong to five orthologous groups (OGs): OG1, 2, 4, 9, and 10 (Wu et al., 2016). OG2 clade genes have the potential to participate in the regulation of reproductive development. *PvSPL6*, *PvSPL7′*, *PvSPL8*, and *PvSPL17* all belong to the OG2 clade. According to the latest research, PvSPL7 and PvSPL8 redundantly regulate flowering in switchgrass. Overexpression of the individual *SPL7* and *SPL8* gene promotes flowering, whereas their individual downregulation moderately delays flowering. Only simultaneous downregulation of *SPL7* and *SPL8* causes significant delayed flowering. PvSPL7 and 8 induce phase transition and flowering in grasses by directly upregulating *SEPALLATA3* (*SEP3*) and *MADS32* (Gou et al., 2019).

Recent studies have shown that the SPLs act as a key hub, integrating various flowering regulation pathways in *Arabidopsis* (Hong and Jackson, 2015; Teotia and Tang, 2015). Photoperiodic and gibberellin pathways have marked effect on the expression of some *SPL* genes. For photoperiodic pathway, the expression of *SPL3*/*4*/*5* is influenced by photoperiod in early vegetative stages (Jung et al., 2012). *FT* as the key component of the systemic flowering signal interacts with *FLOWERING LOCUS D* (*FD*), a meristem-specific bZIP transcription factor, in the shoot apex. FD binds directly to the G-box motifs present in the promoters of *SPL3*/*4*/*5*. Moreover, with respect to the changes in photoperiod, SOC1 binds to the CArG motifs present in the promoters of *SPL3*/*4*/*5* to regulate their expressions. Thus, photoperiod induction can induce SPL gene expression in a *CO*-, *SOC1*-, or *FT*-dependent manner (Teotia and Tang, 2015). For gibberellin pathway, GAs are a group of diterpenoid phytohormones that regulate a variety of events in plant development, including seed germination, stem elongation, leaf expansion, flowering, and fruit development (Sun, 2010; Mcatee et al., 2013; Tuan et al., 2018; Bao et al., 2020). GAs have been shown to regulate these diverse biological processes mainly by overcoming the inhibition of the DELLA proteins, a family of nuclear repressors of the GA response. Because DELLA proteins do not contain canonical DNA-binding domains, they regulate downstream genes by interacting with other TFs (Daviere et al., 2008; Hauvermale et al., 2012; Locascio et al., 2013; Xu et al., 2014). Growing evidence indicates that GA signaling and the miR156/*SPL*s module are connected through direct interactions between DELLAs and SPL TFs. For example, the GA-induced flowering pathway can be integrated into the miR156-mediated flowering pathway through interactions between DELLAs and SPLs. The binding of DELLAs to SPLs has been shown to impair the transcriptional activation of downstream SPL target genes. Consequently, DELLAs delay the floral transition by reducing SPL15-mediated expression of *MADS-box* genes (*SOC1* and *FUL*) in the shoot apex or by repressing the activation of *FT* through inhibition of SPL9 in the leaves (Galvão et al., 2012; Yu et al., 2012; Hyun et al., 2016).

In addition, recent studies have shown that the DELLA-SPL9 module is involved in axillary bud formation. SPL9 inhibits the transcription of the axillary bud identity gene *LAS*, while binding of DELLA to SPL9 attenuates the repression of *LAS* by SPL9, thereby promoting axillary bud initiation (Zhang et al., 2020). However, given the fact that TFs are usually expressed in a tissue-specific and temporally variable manner, questions remain about the contribution of *SPL*s to GA signaling at the tissue or single-cell level. Systematic protein–protein interaction assays and visualization of protein-TF interactions *in vivo* will help us to address this question.

In the past two decades, the proteolysis of membrane-bound TFs (MTFs) has been studied extensively as a novel transcriptional regulatory mechanism (Chen et al., 2008; Seo et al., 2008). MTFs are TFs with transmembrane domains (TMs) that are fixed to membranes in a dormant state. When exposed to developmental and environmental cues, some MTFs undergo proteolytic cleavage, releasing intracellular fragments into the nucleus to control gene transcription (Liu et al., 2008; Seo et al., 2010a,b). As a result, MTFs can rapidly respond to pressures from extracellular or intracellular stimuli (Hoppe et al., 2000). MTFs have been observed in many types of organisms, including plants, animals, and microorganisms (Hoppe et al., 2001; Yang et al., 2014; Xie et al., 2015). Consistent with the activation pathways of more typical TFs, these molecules are delicately regulated at many points throughout the signal transduction process. Cellular stimuli can activate MTF precursors and induce their translocation. Cellular translocation signals include ligand-receptor binding response signals, growth hormones, and many types of stress, including temperature, drought, and salinity (Popovic et al., 2007; Seo et al., 2010a,b; Ma et al., 2012; De Clercq et al., 2013; Misra et al., 2013). Signal transduction in response to stress can enable the visualization of protein–TF interactions. In plants, studies of MTFs have focused on two major TF families, NAM/ATAF1/2/CUC2 (NAC) and basic leucine zipper (bZIP; Chen et al., 2008; Seo et al., 2008). To date, eight *NAC*, three *bZIP*, one *MYB*, and one *PHD* TF have been identified and characterized (Seo, 2014). However, other TF families that contain MTFs have not previously been reported.

In this work, we demonstrate that *PvSPL6*, a miR156-targeted member of the *SPL* family, can regulate flowering time in switchgrass. As the number of same orthologous group with PvSPL7 and PvSPL8, PvSPL6 can independently regulate the flowering time. Inhibition of *PvSPL6* expression causes a markedly delays in flowering. Besides, unlike the homolog AtSPL3/4/5 in *Arabidopsis*, PvSPL6 protein has both nuclear and plasma membranes localization. However, the dual localization of nuclear and plasma membranes only appears in the regeneration stage during switchgrass development process. Exogenous GA_3 application induces the rapid nucleus to plasma membrane translocation of PvSPL6 proteins, and the GA_3 inhibitor (paclobutrazol) application induces the plasma membrane returned to nucleus translocation of PvSPL6 proteins. Hence, *PvSPL6* may be an excellent candidate for genetic modification and improvement of biomass production in bioenergy crops. Furthermore, it is possible to discover a new function and mechanism of PvSPL6 in regulating regeneration by studying how PvSPL6 localization responds to GA pathway.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The wild-type control and transgenic switchgrass plants were generated from a high-quality embryogenic callus line with a single genotype that was obtained by screening a mass of Alamo switchgrass (*P. virgatum* L.) seed. The Alamo switchgrass seed was derived from Noble Research Institute, Ardmore, United States. Embryogenic callus of wild-type control and transgenic plants was cultured in a sterile culture room at 23°C with a 16 h light/8 h dark photoperiod (390 μE/m^2/s) and 80% relative humidity. Wild-type control and transgenic plants were planted in a greenhouse at 26°C under a 16 h light/8 h dark photoperiod (390 μE/m^2/s) and approximately 60% relative humidity. The development of switchgrass plants was divided into five elongation stages (E1–E5) and three reproductive stages (R1–R3) as described previously (Moore et al., 1991; Hardin et al., 2013).

Vector Construction and Plant Transformation

The predicted cDNA sequence of *PvSPL6* (*Pavir.2KG430400*) from the switchgrass genome database v4.1[1] was used to design primers for cloning the full-length coding region and RNAi fragment of *PvSPL6*. About 275 bp fragment representing a non-conserved region in the 5′-UTR and coding sequence of *PvSPL6* was selected as the RNAi fragment. This design can rule out the offtarget effect on other miR156-targeted *SPL* genes. The amplified PCR products were confirmed by Sanger sequencing, respectively. For the overexpression of *PvSPL6*, the correct full-length coding region was inserted into the binary pANIC6B vector by LR recombination reactions (Invitrogen, United States). The pANIC6B vectors contain the *attR1-Cm*^{r}*-ccdB-attR2* cassette for overexpression of the target gene, a plant selectable marker cassette (*hygromycin phosphotransferase*, *hph*), and a visual reporter gene cassette (*GUSPlus*; Mann et al., 2012). For suppression of *PvSPL6*, the verified RNAi fragment was cloned into the RNAi-mediated suppression vector pANIC8B driven by the maize *Ubiquitin* promoter (Mann et al., 2012). The main difference between pANIC6B and pANIC8B is that the pANIC8B vectors contain the *attR1-Cm*^{r}*-ccdB-attR2* cassette downstream of an inverted repeat of itself, resulting in a hairpin loop of the target sequence after recombination and transcription. Then the constructed vectors were transferred into *Agrobacterium tumefaciens* strain EHA105 using the freeze–thaw method (Chen et al., 1994). To generate transgenic plants, the embryogenic callus line with a single genotype was employed for *Agrobacterium*-mediated transformation following the procedure described previously

[1]https://phytozome-next.jgi.doe.gov/

(Xi et al., 2009). The control switchgrass plants were generated by using empty pANIC6B and pANIC8B empty vectors, respectively.

For the construction of vector to observe PvSPL6 subcellular localization, the *GFP* was fused to the C-terminal of *PvSPL6* coding region, and then inserted into the pANIC6B vector by LR recombination reactions. Then verified constructs were transferred into EHA105 and introduced into the embryogenic callus line by *Agrobacterium*-mediated transformation. Hygromycin (Phytotechlab, Lenexa, United States) was used as the selection reagent for the production of $PvSPL6_{OE}$, $PvSPL6_{RNAi}$, and PvSPL6-GFP_{OE} transgenic switchgrass plants. Positive transgenic lines were identified by PCR using specific *hph*, *PvSPL6*, and *PvSPL6*-RNAi primers (**Supplementary Table 1**). The expected sizes of the PCR products were 375, 642, and 275 bp, respectively.

Subcellular Localization Assay

PvSPL6 cDNA fragments encoding the N-terminal membrane spanning domain (amino acid 1–71, cDNA 1–213 bp, *PvSPL6*-N) and the remainder (amino acid 72–214, cDNA 214–642 bp, *PvSPL6*-C) were amplified by PCR, respectively. The full-length and two truncated coding sequences of *PvSPL6* were fused with *GFP* and ligated into the pCambia1300 vector. EHA105 containing the final binary vector pCambia1300::*PvSPL6-GFP*, pCambia1300::*PvSPL6*-N-*GFP*, or pCambia1300::*PvSPL6*-C-*GFP* was injected into leaves of four-week old *Nicotiana benthamiana*. P19 from tomato bushy stunt virus was used to inhibit transgenic silencing (Chen et al., 2007). The resulting fluorescence signal was observed 48–72 h after injection using a FluoView FV1000 confocal laser scanning microscope (Olympus, Japan). The fluorescent dye propidium iodide (PI) was used as a cell plasma membrane marker, the 4′,6-diamidino-2-phenylindole (DAPI) was used as a cell nuclear marker. The primer pairs used for vector construction are listed in **Supplementary Table 1**.

Quantitative Real-Time PCR Analysis

Total RNA was extracted from switchgrass stems using a TRIzol kit (TransGen Biotech, Beijing, China) and was reverse-transcribed into cDNA using a PrimeScript RT Reagent Kit with gDNA Eraser (Takara, Dalian, China) according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was performed in a 20-μl reaction volume that contained 10 μl of SYBR Premix ExTaq (Takara, Dalian, China), 2 μl of cDNA (first strand cDNA, diluted five times), and 0.5 μM of each primer. The primer pairs used for qRT-PCR are listed in **Supplementary Table 1**. *PvUBQ2* (*Pavir.1KG065600*) was used as the reference for normalization (Huang et al., 2014). The cycle thresholds were determined using a Roche Light Cycler 480 II sequence detection system (Roche, Shanghai, China).

Phenotypic Measurements

Flowering time, internode number, internode length, tiller number, and plant height were measured on three biological replicates when plants reached the R1 stage. The I2 internodes were used for the measurement of internode length.

Confocal Laser Scanning Microscopy After Hormone and Plant Growth Regulator

Embryogenic callus were induced from inflorescences of selected three PvSPL6-GFP_{OE} lines. These callus was cultured on SM5 medium {MS0 + 5 mg/L 2,4-D (2,4-Dichlorophenoxyacetic acid) + 0.15 mg/L 6-BA [N-(Phenylmethyl)-9H-purin-6-amine]} supplemented with different hormones and plant growth regulators for 2 weeks. For hormones and growth regulators treatments, different concentrations of 2,4-D (0, 1, 3, and 5 mg/L), 6-BA (0.02, 0.05, 0.1, and 1 mg/L), 6-Furfurylamino-purine (KT; 0, 0.5, 1, and 4 mg/L), gibberellin (GA_3; 0, 10, 100, and 400 mg/L), and paclobutrazol (0, 0.5, 1, and 2 mg/L) were used, respectively. The fluorescence signal of each callus type under each hormone and growth regulator treatment was observed 48–72 h later using a FluoView FV1000 confocal laser scanning microscope (Olympus, Japan).

Statistical Analysis

Three control switchgrass plants, three $PvSPL6_{OE}$ lines, and three $PvSPL6_{RNAi}$ lines were statistical analyzed in this work. The selected transgenic plants were propagated simultaneously with three biological replicates. One-way ANOVA was used for qRT-PCR and phenotypic statistical analysis, and treatment means were separated using Duncan's multiple range test ($p < 0.05$). All the statistical analyses were performed with the SPSS software (IBM SPSS Statistics 25.0, United States).

RESULTS

Molecular Cloning and Sequence Analyses of Switchgrass SBP Transcription Factor *PvSPL6*

Blastn searches against the switchgrass genome (*P. virgatum* v4.1, Phytozome) indicated that these four OG2 genes, *PvSPL6* (*Pavir.2KG430400*), *PvSPL7′* (*Pavir.2NG503700*), *PvSPL8* (*Pavir.2KG430000*), and *PvSPL17* (*Pavir.2NG503500*), were located on chromosome 2. *PvSPL6* and *PvSPL7′* as an allele share over 86% sequence identity, and *PvSPL8* and *PvSPL17* as an allele share 90.7% sequence identity (**Supplementary Figure 1**). Another OG2 gene, *PvSPL7*, has also been reported recently (Gou et al., 2019). We used orthologs of the five PvSPLs in OG2 (PvSPL6, PvSPL7′, PvSPL8, PvSPL17, and PvSPL7) from three genome-sequenced species (*A. thaliana*, *Populus trichocarpa*, and *Oryza sativa*) to construct a phylogenetic tree. The tree showed that PvSPL6 and PvSPL7′ clustered together in a group, implying that they have a close evolutionary relationship and similar functions. By contrast, the distance between PvSPL6 and PvSPL7 on the phylogenetic tree suggested that they may have different functions (**Figure 1A**). The sequence alignment further revealed the variation among *PvSPL6*, *PvSPL7*, and *PvSPL8* as well (**Figure 1B**). These results prompted us to explore PvSPL6 in more detail. Using information from the assembled switchgrass genome database at Phytozome, the full-length sequence of *PvSPL6* was isolated to study its function in switchgrass.

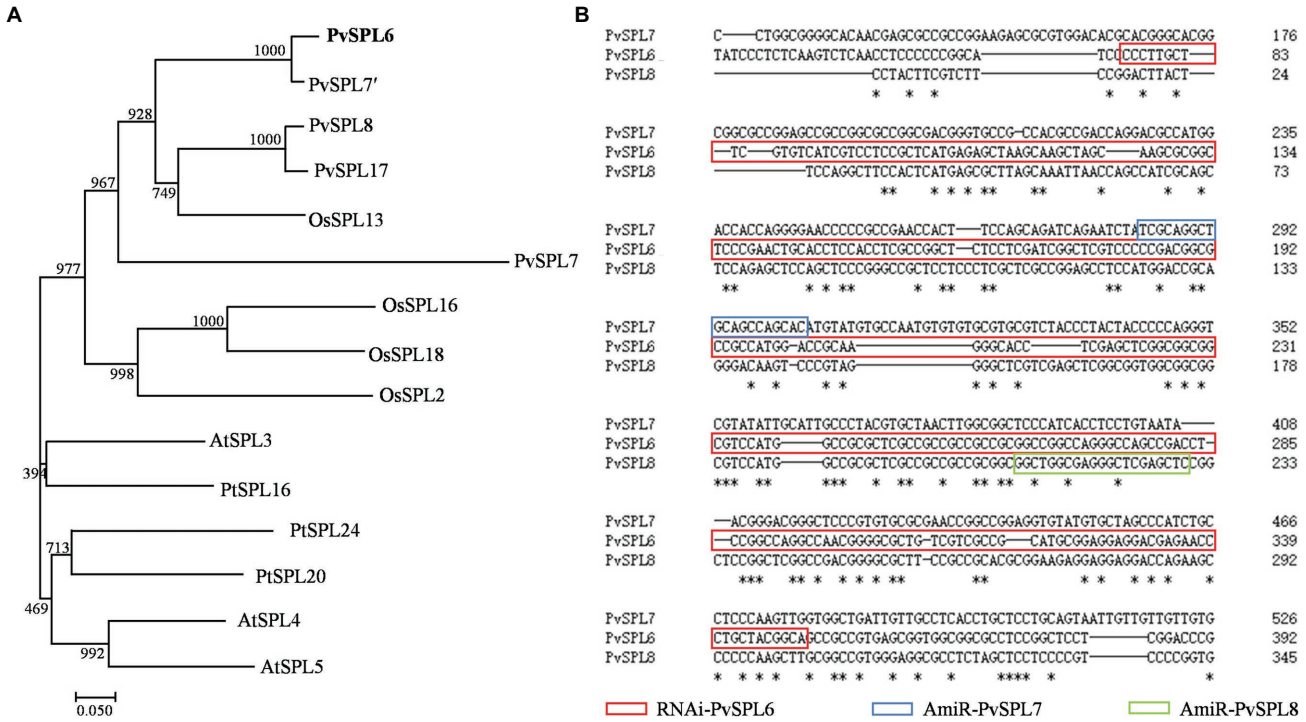


FIGURE 1 | Molecular characterization and sequence analysis of *PvSPL6*. **(A)** Phylogenetic analysis of *SQUAMOSA PROMOTER BINDING-LIKEs* (SPLs) in clade OG2 in dicot and monocot plant species. A maximum likelihood tree was constructed in PhyML version 3.0 based on multiple alignments of deduced protein sequences from *Panicum virgatum* (PvSPL6, Pavir.2KG430400; PvSPL7′, Pavir.2NG503700; PvSPL8, Pavir.2KG430000; and PvSPL17, Pavir.2NG503500), *Oryza sativa* (OsSPL2, LOC_Os01g69830; OsSPL13, LOC_Os07g32170; OsSPL16, LOC_Os08g41940; and OsSPL18, LOC_Os09g32944), *Arabidopsis thaliana* (AtSPL3, AT2G33810; AtSPL4, AT1G53160; and AtSPL5, AT3G15270), and *Populus trichocarpa* (PtSPL16, Potri.011G055900; PtSPL20, Potri.001G398200; and PtSPL24, Potri.007G138800). Bootstrap support values (>50%) based on 1,000 replications are given at the nodes. The sequence data were retrieved from Phytozome and/or Genbank. **(B)** Sequence alignment of *PvSPL6*, *PvSPL7*, and *PvSPL8*. The red box indicates the RNAi regions of *PvSPL6*; the blue box indicates the artificial microRNA (amiRNA) regions of *PvSPL7*; and the green box indicates the amiRNA regions of *PvSPL8*. The artificial microRNA regions mean the nonconserved regions of *PvSPL7* and *8*. The amiRNAs of *PvSPL7* and *8* were to knockdown the expression levels of *PvSPL7* and *8*, respectively. *means the identical nucleotide.

The PvSPL6 TF Shows Both Nuclear and Plasma Membrane Localization

A fused vector containing *PvSPL6* and *GFP* was constructed to investigate the subcellular localization of *PvSPL6*. After Sanger sequencing validation, the pCambia1300::*PvSPL6-GFP* vector was introduced into tobacco leaves by infiltration with *A. tumefaciens* strain EHA105 to produce transient expression. Unlike miR156-targeted PvSPL2 and PvSPL4, the PvSPL6-GFP signal was located in both the nucleus and the plasma membrane (**Figures 2A,B**; **Supplementary Figure 2A**). We then used TMPred[2] to predict TMs in PvSPL6. The results showed that PvSPL6 had one potential transmembrane helix from amino acids 41 to 58 (red label), with a score of 1,035 (only scores above 500 are considered significant; **Figure 2C**). PvSPL2 and PvSPL4 did not contain predicted transmembrane helices (**Supplementary Figure 2B**), consistent with the results of the subcellular localization analysis. To determine the authenticity of TMPred predict, two truncated coding sequences of *PvSPL6*, *PvSPL6*-N (contains the N-terminal membrane spanning domain) and *PvSPL6*-C (remainder) were fused with *GFP* and ligated into the pCambia1300 vector. Transient expression in tobacco leaves showed that the PvSPL6-C-GFP signal was located entirely in the tobacco cell nucleus, whereas the PvSPL6-N-GFP signal was located in both the nucleus and plasma membrane (**Figure 2D**). The SMART tool[3] was used for functional domain analysis. PvSPL6 was shown to have three functional domains, located at amino acids 38–54, amino acids 79–107, and amino acids 110–184, respectively (**Figure 2E**). Among them, the functional domain located at amino acids 110–184 is squamosa promoter binding protein (SBP) domain. It is a highly conserved domain of *SPL* transcription factor family consisting of approximately 78 amino acid residues. Moreover, the coding sequences of the functional domain located at amino acids 38–54 of PvSPL6 was highly coincident with the N-terminal membrane spanning domain. Hence, this functional domain of PvSPL6 has the potential to determine its membrane localization.

[2]https://embnet.vital-it.ch/

[3]http://smart.embl-heidelberg.de/

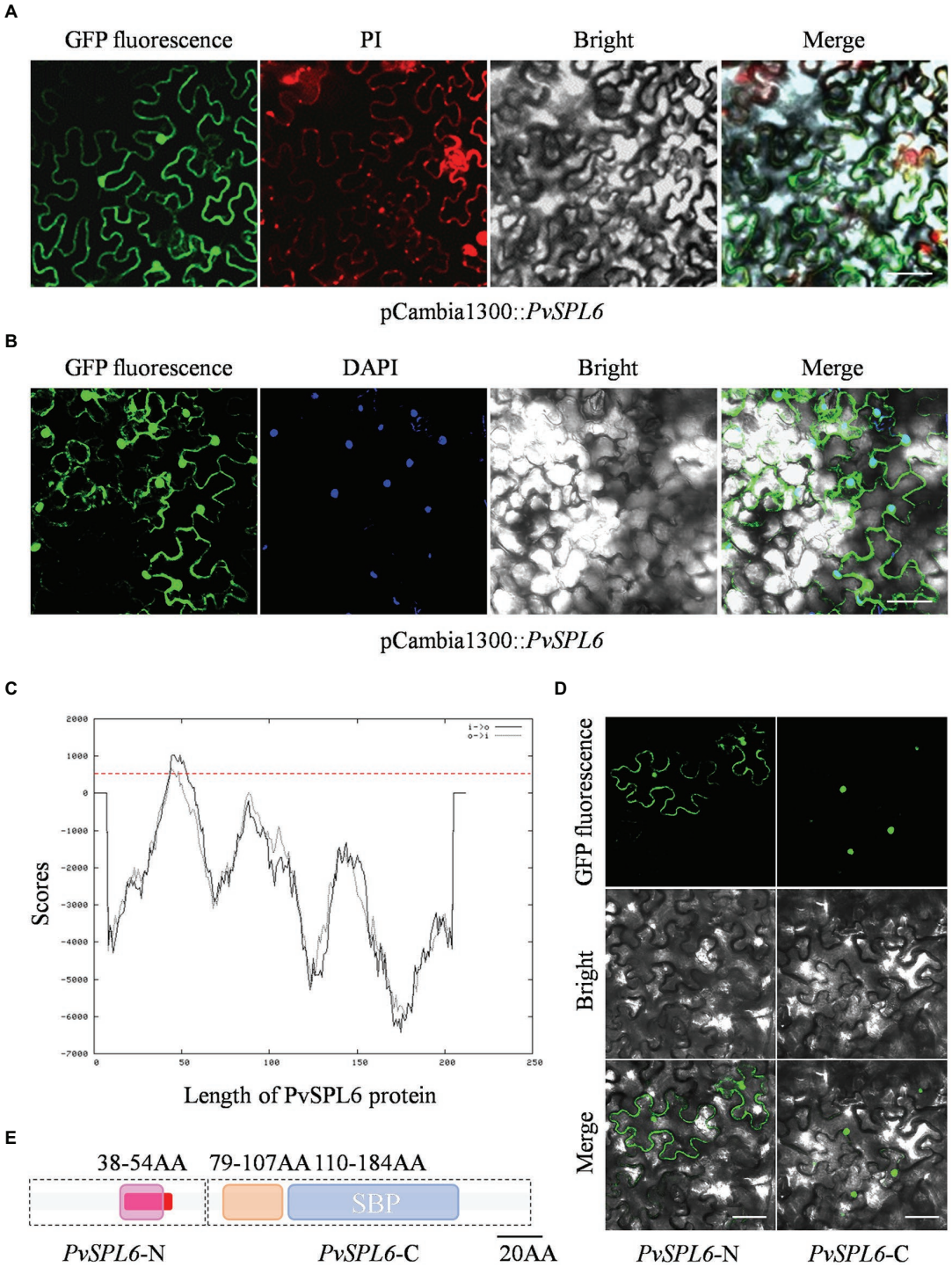


FIGURE 2 | Subcellular localization and bioinformatics analysis of PvSPL6. **(A)** and **(B)** Subcellular localization assays of PvSPL6. *Agrobacterium* cells harboring fusion constructs were infiltrated into the abaxial surfaces of *Nicotiana benthamiana* leaves, and samples were observed 72 h later under a GloMax 20/20 single tube luminometer (Promega, United States). GFP fluorescence, green fluorescent signal; PI, propidium iodide signal; DAPI, 4′,6-diamidino-2-phenylindole signal; Bright, bright field signal; and Merge, superimposed signal. Scale bar = 20 μm. **(C)** Transmembrane domain prediction of PvSPL6 by TMPred. The red dotted line indicates a score of 500 (scores above 500 are considered significant). Black solid line means inside to outside helices; black dotted line means outside to inside helices. **(D)** Subcellular localization assays of PvSPL6-N and PvSPL6-C. GFP fluorescence, green fluorescent signal; Bright, bright field signal; and Merge, superimposed signal. Scale bar = 20 μm. **(E)** Functional domain predictions of PvSPL6. The boxes indicate functional domains. Purple box means the first functional domain with low compositional complexity; orange box means the second functional domain with low compositional complexity; and blue box means conserved SBP domain. The red section indicates the predicted transmembrane region.

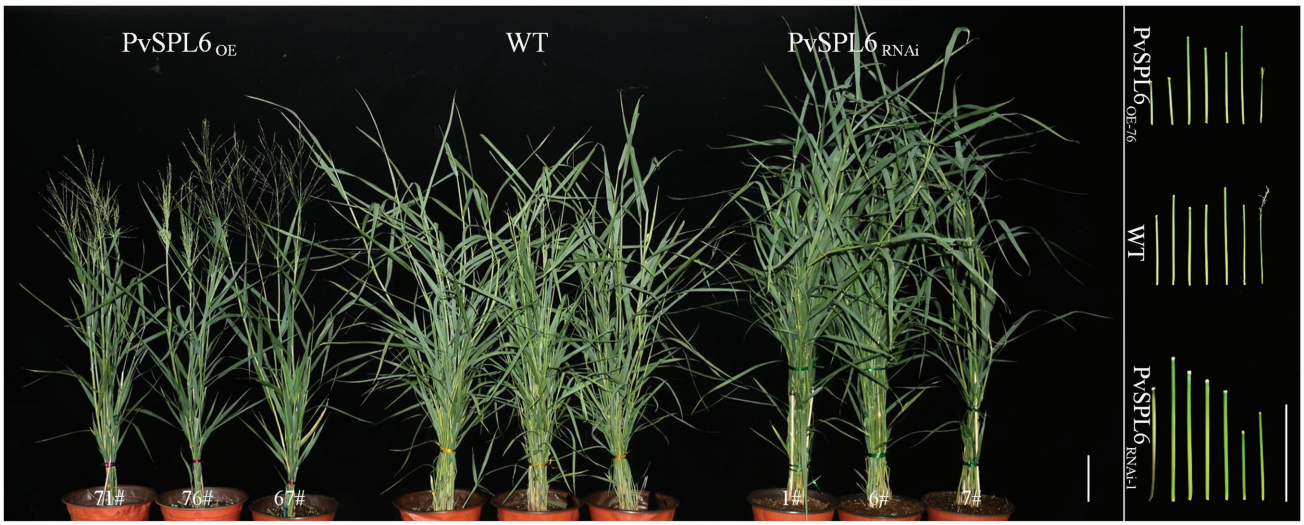


FIGURE 3 | Morphological characterization of *PvSPL6* transgenic switchgrass plants. Morphological characterization of $PvSPL6_{OE}$ and $PvSPL6_{RNAi}$ transgenic switchgrass plants. Scale bar = 10 cm.

Morphological Characterization of *PvSPL6* Transgenic Switchgrass Plants

To characterize the function of *PvSPL6* in switchgrass, we constructed *PvSPL6* overexpression and RNAi vectors and introduced them into wild-type switchgrass callus by *Agrobacterium*-mediated transformation (**Supplementary Figure 3**). Compared with the wild-type control, the $PvSPL6_{OE}$ lines consistently displayed markedly earlier heading dates, reduced internode lengths and numbers, and shorter plant heights. By contrast, the $PvSPL6_{RNAi}$ lines showed conspicuously delayed heading dates and increased internode lengths and numbers (**Figure 3**). We selected the three $PvSPL6_{OE}$ lines with the highest expression levels ($PvSPL6_{OE-67}$, $_{-71}$, and $_{-76}$) and the three $PvSPL6_{RNAi}$ lines with the lowest expression levels ($PvSPL6_{RNAi-1}$, $_{-6}$, and $_{-7}$) for further phenotypic analysis (**Figure 4**; **Supplementary Figure 4**). Taken together, our results showed that *PvSPL6* overexpression and suppression altered plant development. Upregulation of *PvSPL6* shortened the vegetative growth period and decreased dry biomass yield by 40.90, 44.96, and 55.80% in the three lines. Downregulation of *PvSPL6* extended the vegetative growth period and increased the dry biomass yield by 47.73, 44.50, and 62.54% (**Figure 5**). To exclude the possibility that other genes in the same clade were inhibited by *PvSPL6* RNAi, we also measured the expression levels of *PvSPL7′*, *PvSPL7*, *PvSPL8*, and *PvSPL17* in $PvSPL6_{RNAi}$ transgenic plants. Only *PvSPL6* expression was inhibited in $PvSPL6_{RNAi}$ plants relative to the wild-type, confirming that the phenotype of the $PvSPL6_{RNAi}$ plants was caused by reduced expression of *PvSPL6* alone (**Figure 4B**; **Supplementary Figure 5**). The phenotypes of the two transgenic plant types indicated that *PvSPL6* functions in the control the flowering time and affects internode elongation in switchgrass. To explain the observed phenotype, we roughly validated how PvSPL6 participates in regulation of floral transitions. The high correlation between the expression levels of *PvSPL6* and *PvSEP3*/*PvMADS32*, the target genes of PvSPL7 and PvSPL8, in different transgenic lines suggested that PvSPL6 has the similar regulatory mechanism in floral transitions to PvSPL7/8 (**Supplementary Figure 6**).

Subcellular Localization of PvSPL6 in Different Tissues

To further study the localization of PvSPL6, the verified pANIC6B::*PvSPL6-GFP* constructs were transferred into a high-quality embryogenic callus line with a single genotype. By *Agrobacterium*-mediated transformation, we produced PvSPL6-GFP_{OE} transgenic plants. The PvSPL6-GFP_{OE} transgenic plants had phenotypes similar to those of the $PvSPL6_{OE}$ plants. We induced embryogenic callus from three PvSPL6-GFP_{OE} transgenic lines, and we observed PvSPL6-GFP signal only in the nuclei of the undifferentiated transgenic callus (loose and irregular and have not yet formed somatic embryos; **Figure 6A**). This was not consistent with the results, we observed in tobacco leaf cells. However, we observed partial translocation of the PvSPL6-GFP signal from the nucleus to the plasma membrane when the callus was in the differentiation stage (compact and dense somatic embryo, and even appear green bud points; **Figure 6B**). During subsequent development, the differentiated calli formed complete switchgrass plants. Confocal laser scanning microscopy showed that PvSPL6 exhibited its complete nuclear localization in both leaves and roots of the resulting plants (**Figures 6C,D**).

GA_3 Controls the Localization of PvSPL6 in Switchgrass

To investigate the biological significance of PvSPL6 membrane localization, we first needed to identify the factor and related pathway to which PvSPL6 localization responds. Ligand-receptor binding response signals, growth hormones, and many types of stress may be the candidates for influencing

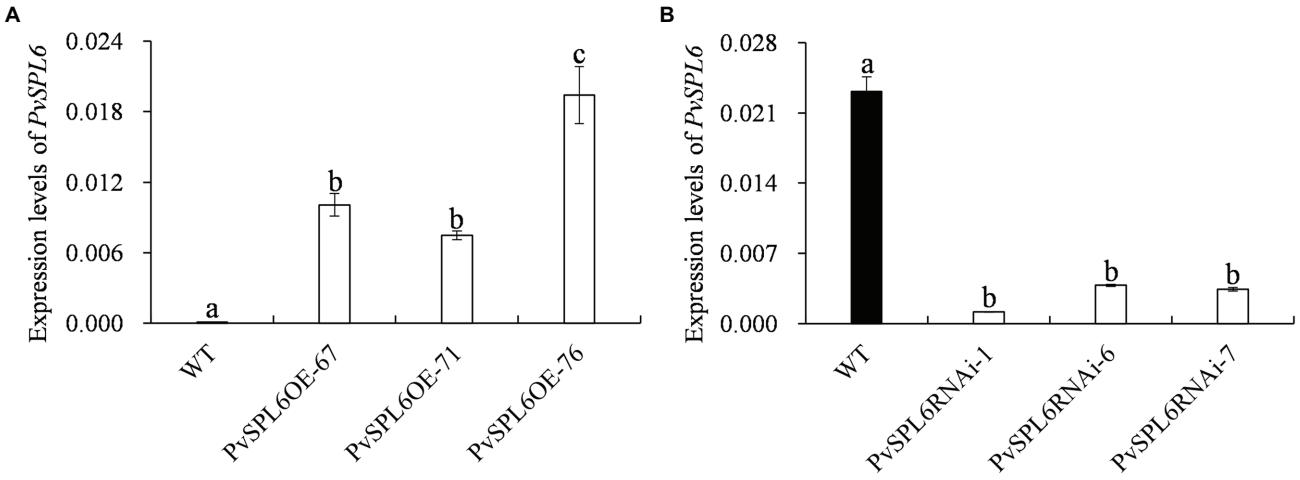


FIGURE 4 | Quantitative real-time PCR (qRT-PCR) analysis of *PvSPL6* transcript levels in different $PvSPL6_{OE}$ and $PvSPL6_{RNAi}$ transgenic switchgrass plants. **(A)** qRT-PCR analysis of *PvSPL6* transcript levels in different $PvSPL6_{OE}$ transgenic switchgrass plants. **(B)** Quantitative real-time PCR analysis of *PvSPL6* transcript levels in different $PvSPL6_{RNAi}$ transgenic switchgrass plants. *PvUBQ2* was used as a reference for normalization. The values are the means ± SEs ($n = 3$). The letters above the error bars indicate significant differences determined by one-way ANOVA ($p < 0.05$, Duncan's multiple-range test).

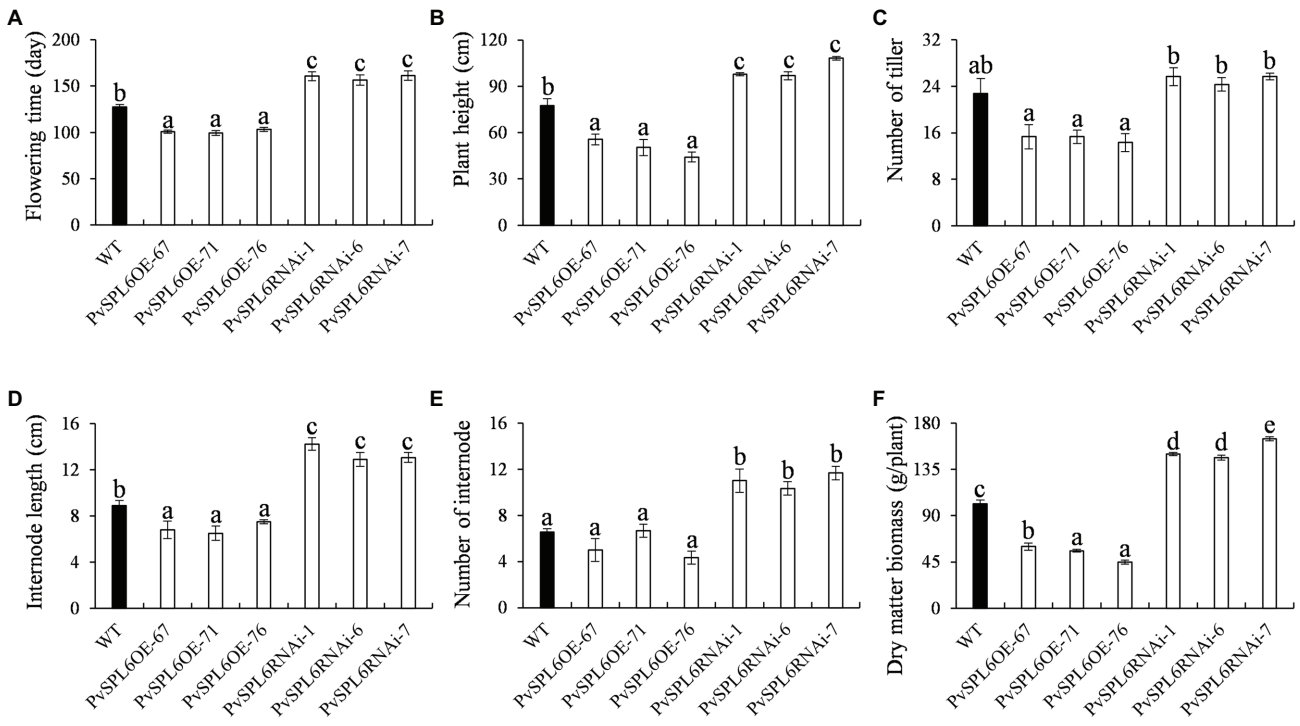


FIGURE 5 | Phenotypic statistics of $PvSPL6_{OE}$ and $PvSPL6_{RNAi}$ transgenic switchgrass plants. Comparisons of **(A)** flowering time, **(B)** plant height, **(C)** number of tiller, **(D)** internode length, **(E)** number of internode, and **(F)** dry matter biomass. The values are the means ± SEs ($n = 3$). The letters above the error bars indicate significant differences determined by one-way ANOVA ($p < 0.05$, Duncan's multiple-range test).

PvSPL6 localization. Combined the flowering phenotype of *PvSPL6* transgenic plants and the roles of GA pathway in flowering regulation, we chose GA_3 to treat $PvSPL6\text{-}GFP_{OE}$ callus. Predictably, we observed partial PvSPL6-GFP signal translocation from the nucleus to the plasma membrane when *PvSPL6* transgenic callus was treated with different concentrations of GA_3 (0, 10, 100 and 400 mg/L). Compared with callus in SM5 medium without GA_3, callus treated with even a low concentration of GA_3 (10 mg/L) showed clear plasma membrane localization of PvSPL6-GFP. This plasma

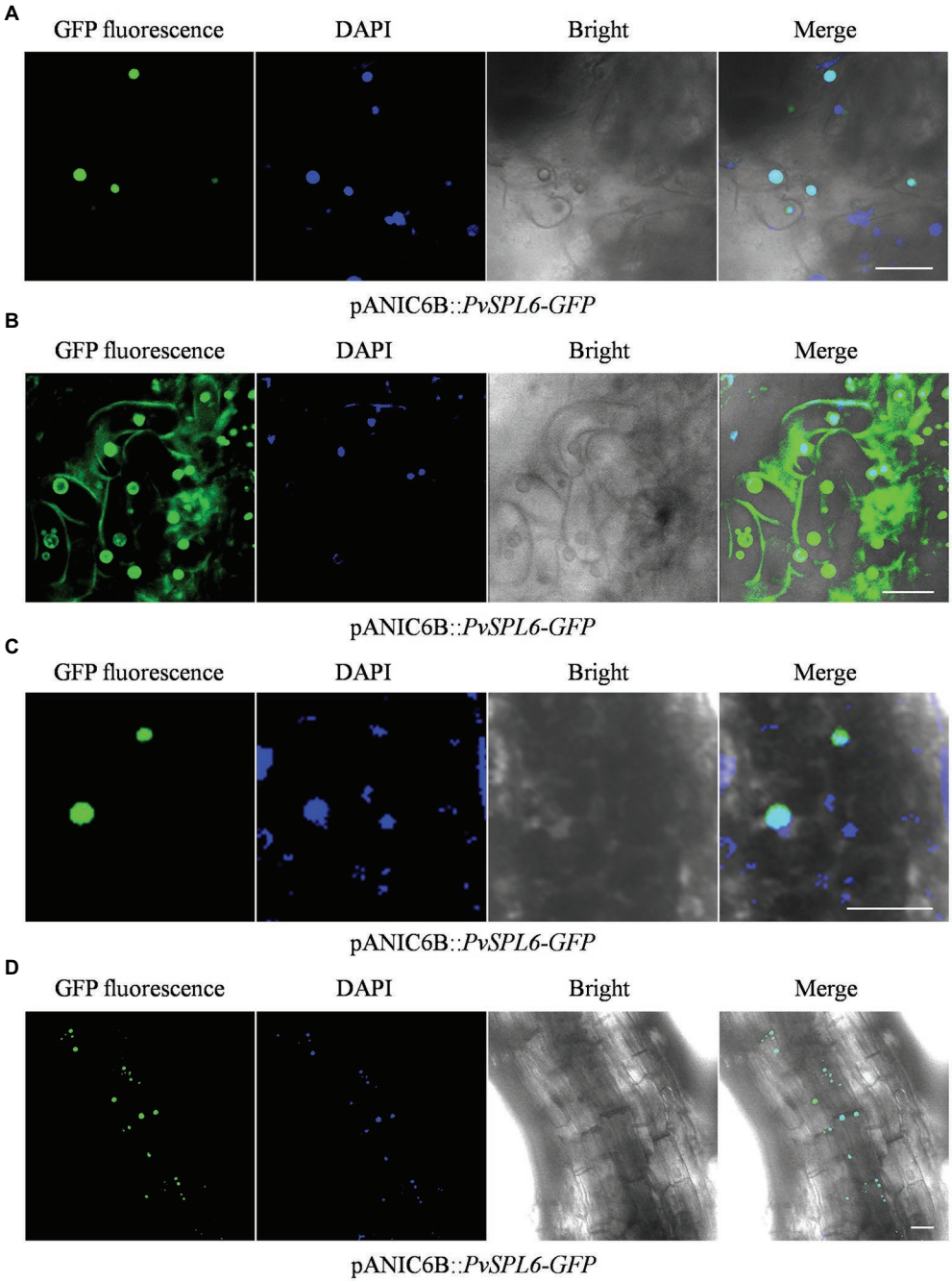


FIGURE 6 | Subcellular localization analysis of different tissues in PvSPL6-GFP$_{OE}$ transgenic plants. **(A)** Subcellular localization analysis of undifferentiated embryogenic callus of PvSPL6-GFP$_{OE}$ transgenic plants. **(B)** Subcellular localization analysis of differentiated embryogenic callus of PvSPL6-GFP$_{OE}$ transgenic plants. **(C)** Subcellular localization analysis of leaves from PvSPL6-GFP$_{OE}$ transgenic plants. **(D)** Subcellular localization analysis of roots from PvSPL6-GFP$_{OE}$ transgenic plants. GFP fluorescence, green fluorescent signal; DAPI, 4′,6-diamidino-2-phenylindole signal; Bright, bright field signal; and Merge, superimposed signal. Scale bar = 20 μm.

membrane localization became more obvious as the GA_3 concentration increased (**Figure 7**). We also used various other plant growth regulators treatments, 2,4-D (0, 1, 3, and 5 mg/L), 6-BA (0.02, 0.05, 0.1, and 1 mg/L), and KT (0, 0.5, 1, and 4 mg/L), to assess the hormone specificity of the PvSPL6 response. Embryogenic callus from three PvSPL6-GFP$_{OE}$ lines was cultured on SM5 medium supplemented with the above compounds at 23°C in the dark for 2 weeks, and the subcellular localization of PvSPL6 was observed by confocal laser scanning microscopy. PvSPL6 maintained its

nuclear localization after treatment with all concentrations of 2,4-D, 6-BA, and KT (**Supplementary Figure 7**). Furthermore, to further confirm the effect of GA_3 on PvSPL6 localization, we cultured differentiated callus with dual localization of nuclear and plasma membrane in medium supplemented with different concentrations of GA_3 inhibitors

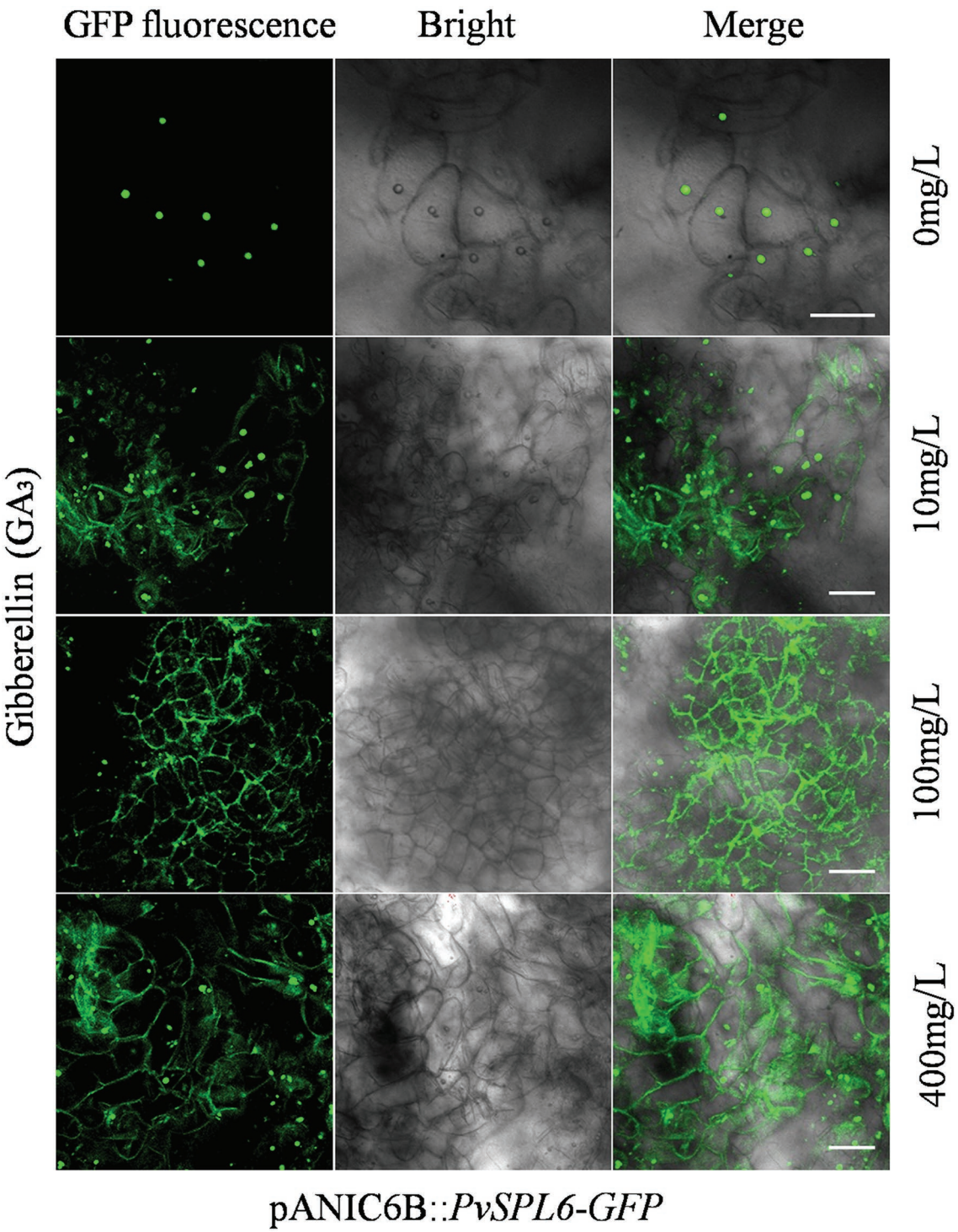


FIGURE 7 | Subcellular localization analysis of embryogenic callus from PvSPL6-GFP_{OE} transgenic plants in response to GA_3. Subcellular localization of PvSPL6 in response to different concentrations of GA_3 (0, 10, 100, and 400 mg/L). GFP fluorescence, green fluorescent signal; Bright, bright field signal; and Merge, superimposed signal. Scale bar = 20 μm.

(paclobutrazol, 0, 0.5, 1, and 2 mg/L). The results showed that PvSPL6-GFP signal translocation from the plasma membrane returned to the nucleus in differentiated callus after paclobutrazol treatment. And this phenomenon became more obvious as the increase of concentration (**Supplementary Figure 8**). In conclusion, the localization of PvSPL6 is more sensitive to plant endogenous hormones compared with plant growth regulators. And GA_3 is the crucial factor responsible for the plasma membrane localization of PvSPL6 in cells.

DISCUSSION

Precise flowering time is critical to reproductive success. Since the discovery that miR156, whose expression decreases with age, mediates the regulation of flowering time in plants, the miR156-*SPL*s module has attracted significant attention as the core regulatory hub of the age pathway. To date, the *SPL* family has been found to promote flowering mainly through three pathways in *Arabidopsis*. SPL3/4/5, SPL9, and SPL2/10/11 are dominant in each pathway, respectively. In contrast to the extensive studies in *Arabidopsis*, little information is available on the flowering-related roles of *SPL*s in the Gramineae. OG2 clade genes have the potential to participate in the regulation of reproductive development in switchgrass. *PvSPL6*, *PvSPL7′*, *PvSPL8*, *PvSPL17*, and *PvSPL7* all belong to the OG2 clade. Based on molecular characteristics and sequence analysis of genes, the subfamily was further divided into three branches: *PvSPL6* and *PvSPL7′*, *PvSPL8* and *PvSPL17*, and *PvSPL7*. Among them, *PvSPL6* and *PvSPL7′*, *PvSPL8* and *PvSPL17* as the allele showed high degree of sequence similarity and close evolutionary relationship. *PvSPL7* showed significant divergence from the *SPL* genes belongs to the same subfamily. Currently, Only *PvSPL7* and *8* have been functionally identified in switchgrass. Overexpression of *PvSPL7* and *8* promotes flowering, whereas downregulation of individual genes moderately delays flowering. Simultaneous downregulation of *PvSPL7* and *8* results in extremely delayed or nonflowering plants (Gou et al., 2019). We therefore studied the function of *PvSPL6* in the present study and found that *PvSPL6* regulated phase transition and flowering in switchgrass. Downregulation of *PvSPL6* by itself significantly delayed flowering, suggesting that *PvSPL6* may be the dominant gene in this subfamily for the regulation of flowering time in switchgrass.

Subcellular localization assays showed that PvSPL6 was localized to both the nucleus and the plasma membrane, unlike its SPL3/4/5 homologs in *Arabidopsis*. Transmembrane domain prediction showed that PvSPL6 contained a transmembrane structure that was not present in AtSPL3/4/5 or in other SPL subfamilies of switchgrass. As an MTF, *PvSPL6* may therefore have unique functions or mechanisms. Large-scale expression profiling of plant MTF genes and phenotypic analyses of available mutants show that MTFs are involved in diverse developmental processes and growth hormone signaling (Kim et al., 2007). The transcriptional control conferred by the activation of plant MTFs is thought to have a wide array of regulatory roles in diverse aspects of plant growth and development. Meanwhile, our data showed that the nuclear and plasma membrane dual localization of PvSPL6 only occurred at the stage of callus differentiation during the whole development process, so we speculated that the plasma membrane localization of PvSPL6 had the potential to participate in switchgrass regeneration. But these speculations need to be further verified. Thus, research on the functional implications of PvSPL6 plasma membrane localization may break new ground and provide additional clues for understanding the molecular mechanisms by which TF activity is regulated.

Furthermore, relevant studies have shown that MTFs mediate diverse aspects of stress response and enable the rapid regulation of transcription under stressful conditions. The rapid turnover of membrane-bound proteins is essential for cell survival, as is the maintenance of a minimum level of physiological activity under stress conditions (Vik and Rine, 2000; Poon and Jans, 2005). Our data indicated that PvSPL6 proteins were rapidly transported from the nucleus to the plasma membrane after exogenous GA_3 application, and returned from the plasma membrane to nucleus after treated with the GA_3 inhibitor (paclobutrazol). Moreover, PvSPL8, the same subfamily with PvSPL6, also showed both nuclear and plasma membrane localization, and responded to GA_3 treatment (**Supplementary Figure 9**). These results confirm that the plasma membrane localization of PvSPL6 subfamily indeed respond to GA_3 signal and have a directly or indirectly related to GA pathway. Previous studies have shown that the binding of DELLAs, components of GA signaling, to SPLs blocks the transcriptional activation of their downstream target genes. DELLAs delay the floral transition by reducing the SPL15-mediated expression of MADS-box genes (*SOC1* and *FUL*) in the shoot apex or by repressing the activation of *FT* in leaves by inhibiting SPL9 (Galvão et al., 2012; Yu et al., 2012; Hyun et al., 2016). In addition, SPL9 represses transcription of the axillary bud identity gene *LATERAL SUPPRESSOR* (*LAS*), and the binding of DELLA to SPL9 attenuates germination (Zhang et al., 2020). We therefore speculate that PvSPL6 may respond to a specific protein in the GA signaling pathway. After receiving this protein signal, the PvSPL6 TF may be activated and then translocated from the nucleus to the membrane, thus curtailing its TF activity in the nucleus. The biological processes and regulatory mechanisms associated with the transportation of PvSPL6 from the nucleus to the plasma membrane in switchgrass are still largely unknown, but it is worth investigating in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

XZ, JC, YW, and CF designed the research. JC, YW, WenL, WeiL, LZ, GC, and YB performed the experiments. JC, YM, XZ, CF, and DM analyzed the data. XZ, JC, and YW wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Key Research and Development Project (2016YFC0501702), Special Projects for The Central Government to guide The Development of Local Science and Technology (2021FRD05023), Ningxia Hui Autonomous Region Key R&D Project (2020BCF01001), Biological Resources Programme, Chinese Academy of Sciences (KFJ-BRP-007-018), and Special Project of Ningxia Academy of Agriculture and Forestry Science and Technology Cooperation with Foreign (DW-X-2020002).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.834431/full#supplementary-material

REFERENCES

Amasino, R. M., and Michaels, S. D. (2010). The timing of flowering. *Plant Physiol.* 154, 516–520. doi: 10.1104/pp.110.161653

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

Bao, S., Hua, C., Shen, L., and Yu, H. (2020). New insights into gibberellin signaling in regulating flowering in *Arabidopsis*. *J. Integr. Plant Biol.* 62, 118–131. doi: 10.1111/jipb.12892

Campos-Rivero, G., Osorio-Montalvo, P., Sánchez-Borges, R., Us-Camas, R., Duarte-Aké, F., and De-la-Peña, C. (2017). Plant hormone signaling in flowering: an epigenetic point of view. *J. Plant Physiol.* 214, 16–27. doi: 10.1016/j.jplph.2017.03.018

Casler, M. D. (2012). *Switchgrass Breeding, Genetics, and Genomics*. London: Springer. 29–53.

Chen, W. X., Chen, J., Zhang, Z. Z., and Huang, A. L. (2007). P19 of tomato bushy stunt virus suppresses RNA silencing induced by short hairpin RNA in mammal cells. *Virol. Sin.* 22, 199–206. doi: 10.1007/s12250-007-0022-3

Chen, H., Nelson, R. S., and Sherwood, J. L. (1994). Enhanced recovery of transformants of *Agrobacterium tumefaciens* after freeze-thaw transformation and drug selection. *Biotechniques* 16, 664–668. doi: 10.1006/abio.1994.1170

Chen, Y. N., Slabaugh, E., and Brandizzi, F. (2008). Membrane-tethered transcription factors in *Arabidopsis thaliana*: novel regulators in stress response and development. *Curr. Opin. Plant Biol.* 11, 695–701. doi: 10.1016/j.pbi.2008.10.005

Cho, L. H., Yoon, J., and An, G. (2017). The control of flowering time by environmental factors. *Plant J.* 90, 708–719. doi: 10.1111/tpj.13461

Daviere, J. M., de, L. M., and Prat, S. (2008). Transcriptional factor interaction: a central step in DELLA function. *Curr. Opin. Genet. Dev.* 18, 295–303. doi: 10.1016/j.gde.2008.05.004

De Clercq, I., Vermeirssen, V., Van Aken, O., Vandepoele, K., Murcha, M. W., Law, S. R., et al. (2013). The membrane-bound NAC transcription factor ANAC013 functions in mitochondrial retrograde regulation of the oxidative stress response in *Arabidopsis*. *Plant Cell* 25, 3472–3490. doi: 10.1105/tpc.113.117168

Fu, C. X., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M. J., et al. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc. Natl. Acad. Sci. U. S. A.* 108, 3803–3808. doi: 10.1073/pnas.1100310108

Fu, C. X., Sunkar, R., Zhou, C. E., Shen, H., Zhang, J. Y., Matts, J., et al. (2012). Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol. J.* 10, 443–452. doi: 10.1111/j.1467-7652.2011.00677.x

Galvão, V. C., Horrer, D., Küttner, F., and Schmid, M. (2012). Spatial control of flowering by DELLA proteins in *Arabidopsis thaliana*. *Development* 139, 4072–4082. doi: 10.1242/dev.080879

Gou, J. Q., Tang, C. R., Chen, N. C., Wang, H., Debnath, S., Sun, L., et al. (2019). *SPL7* and *SPL8* represent a novel flowering regulation mechanism in switchgrass. *New Phytol.* 222, 1610–1623. doi: 10.1111/nph.15712

Hardin, C. F., Fu, C. X., Hisano, H., Xiao, X., Shen, H., Stewart, C. N., et al. (2013). Standardization of switchgrass sample collection for cell wall and biomass trait analysis. *Bioenerg. Res.* 6, 755–762. doi: 10.1007/s12155-012-9292-1

Hauvermale, A. L., Ariizumi, T., and Steber, C. M. (2012). Gibberellin signaling: a theme and variations on DELLA repression. *Plant Physiol.* 160, 83–92. doi: 10.1104/pp.112.200956

Hill, C. B., and Li, C. (2016). Genetic architecture of flowering phenology in cereals and opportunities for crop improvement. *Front. Plant Sci.* 7:1906. doi: 10.3389/fpls.2016.01906

Hong, Y., and Jackson, S. (2015). Floral induction and flower formation--the role and potential applications of miRNAs. *Plant Biotechnol. J.* 13, 282–292. doi: 10.1111/pbi.12340

Hoppe, T., Kai, M., Rape, M., Schlenker, S., Ulrich, H. D., and Jentsch, S. (2000). Activation of a membrane-bound transcription factor by regulated ubiquitin/proteasome-dependent processing. *Cell* 102, 577–586. doi: 10.1016/s0092-8674(00)00080-5

Hoppe, T., Rape, M., and Jentsch, S. (2001). Membrane-bound transcription factors: regulated release by RIP or RUP. *Curr. Opin. Cell Biol.* 13, 344–348. doi: 10.1016/s0955-0674(00)00218-0

Huang, L. K., Yan, H. D., Jiang, X. M., Zhang, X. Q., Zhang, Y. W., Huang, X., et al. (2014). Evaluation of candidate reference genes for normalization of quantitative RT-PCR in switchgrass under various abiotic stress conditions. *Bioenergy Res.* 7, 1201–1211. doi: 10.1007/s12155-014-9457-1

Hyun, Y., Richter, R., Vincent, C., Martinez-Gallegos, R., Porri, A., and Coupland, G. (2016). Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the *Arabidopsis* shoot meristem. *Dev. Cell* 37, 254–266. doi: 10.1016/j.devcel.2016.04.001

Inoescu, I. A., Møller, B. L., and Sácchez-Párez, R. (2017). Chemical control of flowering time. *J. Exp. Bot.* 68, 369–382. doi: 10.1093/jxb/erw427

Johnson, C. R., Millwood, R. J., Tang, Y., Gou, J., Sykes, R. W., Turner, G. B., et al. (2017). Field-grown miR156 transgenic switchgrass reproduction, yield, global gene expression analysis, and bioconfinement. *Biotechnol. Biofuels* 10:255. doi: 10.1186/s13068-017-0939-1

Jung, J. H., Ju, Y., Seo, P. J., Lee, J. H., and Park, C. M. (2012). The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in *Arabidopsis*. *Plant J.* 69, 577–588. doi: 10.1111/j.1365-313X.2011.04813.x

Jung, J. H., Seo, Y. H., Seo, P. J., Reyes, J. L., Yun, J., Chua, N. H., et al. (2007). The *GIGANTEA*-regulated microRNA172 mediates photoperiodic flowering independent of *CONSTANS* in *Arabidopsis*. *Plant Cell* 19, 2736–2748. doi: 10.1105/tpc.107.054528

Kim, S. Y., Kim, S. G., Kim, Y. S., Seo, P. J., Bae, M., Yoon, H. K., et al. (2007). Exploring membrane-associated NAC transcription factors in *Arabidopsis*: implications for membrane biology in genome regulation. *Nucleic Acids Res.* 35, 203–213. doi: 10.1093/nar/gkl1068

Klein, J., Saedler, H., and Huijser, P. (1996). A new family of DNA binding proteins includes putative transcriptional regulators of the *Antirrhinum majus* floral meristem identity gene *SQUAMOSA*. *Mol. Gen. Genet.* 250, 7–16. doi: 10.1007/BF02191820

Litt, A., and Irish, V. F. (2003). Duplication and diversification in the *APETALA1/FRUITFULL* floral homeotic gene lineage: implications for the evolution of floral development. *Genetics* 165, 821–833. doi: 10.1093/genetics/165.2.821

Liu, J. X., Srivastava, R., Che, P., and Howell, S. H. (2008). Salt stress signaling in *Arabidopsis thaliana* involves a membrane-bound transcription factor AtbZIP17 as a signal transducer. *Plant Signal. Behav.* 3, 56–57. doi: 10.4161/psb.3.1.4889

Locascio, A., Blázquez, M. A., and Alabadí, D. (2013). Genomic analysis of DELLA protein activity. *Plant Cell Physiol.* 54, 1229–1237. doi: 10.1093/pcp/pct082

Ma, Z., Que, H., Ni, Y., Huang, H., Liu, Y., Liu, T., et al. (2012). Cloning and characterization of SCIRR69: a novel transcriptional factor belonging to the CREB/ATF family. *Mol. Biol. Rep.* 39, 7665–7672. doi: 10.1007/s11033-012-1601-4

Mann, D. G., Lafayette, P. R., Abercrombie, L. L., King, Z. R., Mazarei, M., Halter, M. C., et al. (2012). Gateway-compatible vectors for high-throughput gene functional analysis in switchgrass (*Panicum virgatum* L.) and other monocot species. *Plant Biotechnol. J.* 10, 226–236. doi: 10.1111/j.1467-7652.2011.00658.x

Mathieu, J., Warthmann, N., Küttner, F., and Schmid, M. (2007). Export of FT protein from phloem companion cells is sufficient for floral induction in *Arabidopsis*. *Curr. Biol.* 17, 1055–1060. doi: 10.1016/j.cub.2007.05.009

Mcatee, P., Karim, S., Schaffer, R., and David, K. (2013). A dynamic interplay between phytohormones is required for fruit development, maturation, and ripening. *Front. Plant Sci.* 4:79. doi: 10.3389/fpls.2013.00079

Misra, J., Kim, D. K., Choi, W., Koo, S. H., Lee, C. H., Back, S. H., et al. (2013). Transcriptional cross talk between orphan nuclear receptor ERRγ and transmembrane transcription factor ATF6α coordinates endoplasmic reticulum stress response. *Nucleic Acids Res.* 41, 6960–6974. doi: 10.1093/nar/gkt429

Moore, K. J., Moser, L. E., Vogel, K. P., Waller, S. S., Johnson, B. E., and Pedersen, J. F. (1991). Describing and quantifying growth stages of perennial forage grasses. *Agron. J.* 83, 1073–1077. doi: 10.2134/agronj1991.00021962008300060029x

Park, H. J., Kim, W. Y., Pardo, J. M., and Yun, D. J. (2016). Molecular interactions between flowering time and abiotic stress pathways. *Int. Rev. Cell Mol. Biol.* 327, 371–412. doi: 10.1016/bs.ircmb.2016.07.001

Poethig, R. S. (2009). Small RNAs and developmental timing in plants. *Curr. Opin. Genet. Dev.* 19, 374–378. doi: 10.1016/j.gde.2009.06.001

Poon, I. K., and Jans, D. A. (2005). Regulation of nuclear transport: central role in development and transformation? *Traffic* 6, 173–186. doi: 10.1111/j.1600-0854.2005.00268.x

Popovic, M., De Biasio, A., Pintar, A., and Pongor, S. (2007). The intracellular region of the notch ligand Jagged-1 gains partial structure upon binding to synthetic membranes. *FEBS J.* 274, 5325–5336. doi: 10.1111/j.1742-4658.2007.06053.x

Seo, P. J. (2014). Recent advances in plant membrane-bound transcription factor research: emphasis on intracellular movement. *J. Integr. Plant Biol.* 56, 334–342. doi: 10.1111/jipb.12139

Seo, P. J., Kim, S. G., and Park, C. M. (2008). Membrane-bound transcription factors in plants. *Trends Plant Sci.* 13, 550–556. doi: 10.1016/j.tplants.2008.06.008

Seo, P. J., Kim, M. J., Park, J. Y., Kim, S. Y., Jeon, J., Lee, Y. H., et al. (2010a). Cold activation of a plasma membrane-tethered NAC transcription factor induces a pathogen resistance response in *Arabidopsis*. *Plant J.* 61, 661–671. doi: 10.1111/j.1365-313X.2009.04091.x

Seo, P. J., Kim, M. J., Song, J. S., Kim, Y. S., Kim, H. J., and Park, C. M. (2010b). Proteolytic processing of an *Arabidopsis* membrane-bound NAC transcription factor is triggered by cold-induced changes in membrane fluidity. *Biochem. J.* 427, 359–367. doi: 10.1042/BJ20091762

Sun, T. P. (2010). Gibberellin Signal Transduction in Stem Elongation & Leaf Growth. *Plant Hormones* 15, 308–328. doi: 10.1007/978-1-4020-2686-7_15

Teotia, S., and Tang, G. (2015). To bloom or not to bloom: role of microRNAs in plant flowering. *Mol. Plant* 8, 359–377. doi: 10.1016/j.molp.2014.12.018

Tuan, P. A., Kumar, R., Rehal, P. K., Toora, P. K., and Ayele, B. T. (2018). Molecular mechanisms underlying abscisic acid/gibberellin balance in the control of seed dormancy and germination in cereals. *Front. Plant Sci.* 9:668. doi: 10.3389/fpls.2018.00668

Vik, A., and Rine, J. (2000). Membrane biology: membrane-regulated transcription. *Curr. Biol.* 10, R869–R871. doi: 10.1016/s0960-9822(00)00822-8

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, 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, Z. Y., and Ge, Y. X. (2006). Invited review: recent advances in genetic transformation of forage and turf grasses. *In Vitro Cell Dev. Biol. Plant* 42, 1–18. doi: 10.1079/IVP2005726

Wu, Z. Y., Cao, Y. P., Yang, R. J., Qi, T. X., Hang, Y. Q., Lin, H., et al. (2016). Switchgrass *SBP-box* transcription factors *PvSPL1* and *2* function redundantly to initiate side tillers and affect biomass yield of energy crop. *Biotechnol. Biofuels* 9:101. doi: 10.1186/s13068-016-0516-z

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

Xi, Y. J., Fu, C. X., Ge, Y. X., Nandakumar, R., Hisano, H., Bouton, J., et al. (2009). *Agrobacterium*-mediated transformation of switchgrass and inheritance of the transgenes. *Bioenergy Res.* 2, 275–283. doi: 10.1007/s12155-009-9049-7

Xie, J., Han, F., and Shi, Y. (2015). Single-prolonged stress activates the transcription factor ATF6α branch of the unfolded protein response in rat neurons of dorsal raphe nucleus. *Mol. Cell. Biochem.* 399, 209–216. doi: 10.1007/s11010-014-2247-4

Xu, M. L., Hu, T. Q., Zhao, J. F., Park, M. Y., Earley, K. W., Wu, G., et al. (2016). Developmental functions of miR156-regulated *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes in *Arabidopsis thaliana*. *PLoS Genet.* 12:e1006263. doi: 10.1371/journal.pgen.1006263

Xu, H., Liu, Q., Yao, T., and Fu, X. (2014). Shedding light on integrative GA signaling. *Curr. Opin. Plant Biol.* 21, 89–95. doi: 10.1016/j.pbi.2014.06.010

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

Yang, Z., Wang, X., Gu, S., Hu, Z., Xu, H., and Xu, C. (2007). Comparative study of SBP-box gene family in *Arabidopsis* and rice. *Gene* 407, 1–11. doi: 10.1016/j.gene.2007.02.034

Yang, Z. T., Wang, M. J., Sun, L., Lu, S. J., Bi, D. L., Sun, L., et al. (2014). The membrane-associated transcription factor NAC089 controls ER-stress-induced programmed cell death in plants. *PLoS Genet.* 10:e1004243. doi: 10.1371/journal.pgen.1004243

Yu, S., Galvao, 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, Q. Q., Wang, J. G., Wang, L. Y., Wang, J. F., Wang, Q., Yu, P., et al. (2020). Gibberellin repression of axillary bud formation in *Arabidopsis* by modulation of DELLA-SPL9 complex activity. *J. Integr. Plant Biol.* 62, 421–432. doi: 10.1111/jipb.12818

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 February 2022
doi: 10.3389/fpls.2022.819188

Identification of Chilling Accumulation-Associated Genes for Litchi Flowering by Transcriptome-Based Genome-Wide Association Studies

Xingyu Lu[1,2†], Peitao Lü[3†], Hao Liu[1†], Houbin Chen[1], Xifen Pan[1], Pengxu Liu[1], Lei Feng[1], Silin Zhong[4] and Biyan Zhou[1*]*

[1] Guangdong Litchi Engineering Research Center, College of Horticulture, South China Agricultural University, Guangzhou, China, [2] College of Life and Health Science, Kaili University, Kaili, China, [3] College of Horticulture, Fujian Agriculture and Forestry University-University of California Riverside (FAFU-UCR) Joint Center for Horticultural Biology and Metabolomics, Haixia Institute of Science and Technology, Fujian Agriculture and Forestry University, Fuzhou, China, [4] State Key Laboratory of Agrobiotechnology, School of Life Sciences, Chinese University of Hong Kong, Shatin, Hong Kong SAR, China

Litchi is an important Sapindaceae fruit tree. Flowering in litchi is triggered by low temperatures in autumn and winter. It can be divided into early-, medium-, and late-flowering phenotypes according to the time for floral induction. Early-flowering varieties need low chilling accumulation level for floral induction, whereas the late-flowering varieties require high chilling accumulation level. In the present study, RNA-Seq of 87 accessions was performed and transcriptome-based genome-wide association studies (GWAS) was used to identify candidate genes involved in chilling accumulation underlying the time for floral induction. A total of 98,155 high-quality single-nucleotide polymorphism (SNP) sites were obtained. A total of 1,411 significantly associated SNPs and 1,115 associated genes (AGs) were identified, of which 31 were flowering-related, 23 were hormone synthesis-related, and 27 were hormone signal transduction-related. Association analysis between the gene expression of the AGs and the flowering phenotypic data was carried out, and differentially expressed genes (DEGs) in a temperature-controlled experiment were obtained. As a result, 15 flowering-related candidate AGs (CAGs), 13 hormone synthesis-related CAGs, and 11 hormone signal transduction-related CAGs were further screened. The expression levels of the CAGs in the early-flowering accessions were different from those in the late-flowering ones, and also between the flowering trees and non-flowering trees. In a gradient chilling treatment, flowering rates of the trees and the CAGs expression were affected by the treatment. Our present work for the first time provided candidate genes for genetic regulation of flowering in litchi using transcriptome-based GWAS.

Keywords: flowering, chilling accumulation, transcriptome, association analysis, litchi

OPEN ACCESS

Edited by:
Yuexue Liu,
Shenyang Agricultural University, China

Reviewed by:
Shengnan Huang,
Shenyang Agricultural University, China
Pedro Martinez-Gomez,
Center for Edaphology and Applied Biology of Segura, Spanish National Research Council (CSIC), Spain

***Correspondence:**
Silin Zhong
silin.zhong@cuhk.edu.hk
Biyan Zhou
zhoubiyan@scau.edu.cn

† These authors have contributed equally to this work

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

Received: *21 November 2021*
Accepted: *24 January 2022*
Published: *23 February 2022*

Citation:
Lu X, Lü P, Liu H, Chen H, Pan X, Liu P, Feng L, Zhong S and Zhou B (2022) Identification of Chilling Accumulation-Associated Genes for Litchi Flowering by Transcriptome-Based Genome-Wide Association Studies. Front. Plant Sci. 13:819188. doi: 10.3389/fpls.2022.819188

INTRODUCTION

Litchi is an important Sapindaceae fruit tree. Flowering in litchi is triggered by low temperatures in autumn and winter (Menzel and Simpson, 1988; Chen and Huang, 2005; Zhou et al., 2014). It is widely grown in southeast Asia within the latitude of 17°–26°N, and is commercially cultivated in South Africa and Australia within the latitude of 17°–32°S. These areas are normally chill in winter without freezing temperatures, in which chilling accumulation is enough for litchi floral induction. However, during the past 20 years, insufficient winter chilling attributed to global warming that has been occurring frequently. An extreme case happened in the winter of 2018–2019, in which the monthly average temperature was 17.8°C from November to January, resulting in poor flowering and low yield in 2019 (Su et al., 2020). Hence, it is important to understand the genetics of litchi flowering in relation to chilling so as to find ways for its regulation.

As the origin center of litchi, China is rich in litchi resources with diverse traits. One of the most important traits of these resources is the time for the accomplishment of floral induction, which is related to the chilling requirement. Up till now, all the litchi accessions found are evergreen fruit trees. They are grown in areas with chill winter without freezing temperatures. Unlike the deciduous fruit trees such as almond, whose leaves abscise and floral buds are at the endodormancy stage in winter (Prudencio et al., 2018, 2021), litchi floral buds are induced in winter, and winter chilling is indispensable for floral induction (Menzel and Simpson, 1988; Chen and Huang, 2005; Zhou et al., 2014). In the model plant Arabidopsis, CONSTANS (CO) is regarded as a direct activator of *FLOWERING LOCUS T* (*FT*), and the FT protein is a long-seeking florigen that migrated from leaves to the shoot apical meristem (SAM) to promote floral initiation (Corbesier et al., 2007; Yang et al., 2007). Interestingly, litchi trees grown under high-temperature conditions with only a few leaves treated with low temperature could still produce flowers (Zhang et al., 2017), suggesting the chilled leaves produced signals and migrated from leaves to the SAM to induce floral transition. Litchi resources can be divided into early-, medium-, and late-flowering groups according to the time for floral induction and blossom. The early-flowering group can undergo floral induction early, just at the very beginning of winter, suggesting that it needs a low chilling accumulation level, while the late-flowering group has to be subjected to a long period of winter chilling, suggesting that it needs a high chilling accumulation level. We have recorded the time for floral initiation of the collected accessions in our litchi germplasm garden and investigated the chilling accumulation requirement for floral induction of the accessions as basic information for years.

In the present study, we selected 87 accessions from our litchi germplasm garden and recorded the time for floral initiation, quantified the time and the chilling accumulation into quantitative traits as the flowering phenotype data. We then performed RNA-Seq of the accessions and collected population single-nucleotide polymorphism (SNP) data for transcriptome-based GWAS to identify associated genes (AGs). We also performed RNA-Seq of leaves in flowering and non-flowering trees. Based on this conjoint analysis of RNA-Seq, we narrowed the AGs and identified candidate AGs (CAGs). At last, a gradient chilling treatment of the litchi trees was carried out and the expressions of the flowering-related CAGs were further determined by quantitative real-time RT-PCR (qRT-PCR) to identify crucial genes of flowering regulation underlying chilling accumulation. Our present work for the first time provided CAGs by transcriptome-based GWAS in litchi. These genes may be applied for genetic regulation of flowering, and for bypassing or partly bypassing chilling underlying climate change and global warming.

MATERIALS AND METHODS

Plant Material and Experimental Procedures

For the transcriptome-based GWAS, 87 litchi accessions (cultivars or lines) were selected from the germplasm resource garden in South China Agricultural University, Guangzhou, China (latitude 23°9040″ N, longitude 113°21018″ E). The origin, the pedigree, and the flowering property of the accessions are shown in **Supplementary Table 1**. Leaves of the accessions were collected in early December when temperature was low enough for floral induction. Twenty mature leaves collected from four directions of the terminal shoots from each accession were pooled together as one sample. The samples were immediately frozen in liquid nitrogen and stored at −80°C for RNA extraction.

To compare the gene expression pattern of the flowering and non-flowering litchi trees, 20 5-year-old air-laying potted trees (*Litchi chinensis* cv. Nuomici, 1–1.5 m height) with similar phenological stages were used for controlled temperature treatment. All the trees were grown in 30-L pots with loam, mushroom cinder, and coconut chaff (v: v: v, 3:1:1). When the terminal shoots matured, 10 trees as the non-flowering control (H, high-temperature) were transferred to a growth chamber at 25/20°C (day/night temperature, 12 h day and 12 h night) with a relative humidity of 75–85% and natural light. The remaining 10 trees, as the flowering treatment (L, low-temperature), were transferred to a growth chamber at 15/8°C with the same light and humidity conditions. The trees were rewarmed to 25/20°C after 60 days of chilling. Leaves of the L- and H-treated trees at the time points of 3, 30, 60, and 75 days (floral initiation) were sampled, and leaves of the terminal shoots from 3 to 4 trees were pooled together as a replicate (each time point with three biological replicates). For the RNA-Seq data confirmation, litchi potted trees were grown in chambers with similar temperature and light conditions. Leaves of the L- and H-treated trees at the time points of 0, 30, 60, and 75 days (floral initiation) were sampled for analysis. The samples were immediately frozen in liquid nitrogen and stored at −80°C for RNA extraction and RNA-Seq.

To further confirm the chilling accumulation-related candidate genes, the medium-late-flowering 'Guiwei' litchi trees were selected. A total of 15 5-year-old air-laying potted trees with similar phenological stages were grown in a growth chamber at 10°C for floral induction. The trees were evenly divided into

three groups and transferred to a growth chamber at 22°C after 5 weeks (Treatment 1), 7 weeks (Treatment 2), and 9 weeks (treatment 3, sufficient chilling accumulation), respectively. Leaves of the trees at the time points of 0 day (S1), 21 days (S2), and 35 days (S3) for Treatment 1, S1, S2, S3, and 49 days (S4) for Treatment 2, S1, S2, S3, S4, and 63 days (S5) for Treatment 3 were sampled. All the leaf samples were immediately frozen in liquid nitrogen and stored at −80°C for RNA extraction.

Data Processing of Flowering Phenotype of the Accessions

For trait surveys, we recorded the date of the emergence of panicle primordia visible as "whitish millet." We also recorded the temperatures every 30 min from 1 September 2014 to 30 February 2015 using a temperature and humidity recorder (ZDR-M20). To quantify the specific flowering time into quantitative traits, we used three sets of data to describe the flowering phenotypes. Symbol A, flowering phenotype based on the time of panicle primordium ("whitish millet") emergence, included eight groups. Groups 1–2, 3, 4–5, 6, 7–8 indicated the early-, the early-medium-, the medium-, the medium-late, and the late flowering accessions, respectively. Symbol B indicated flowering phenotype based on the days required for floral induction (from 1 September 2014 to the date when the "whitish millet" appeared). Symbol C indicated flowering phenotype based on chilling accumulation denoted by degree hours according to the method described by Chen et al. (2016), as the sum of the temperatures lower than 20°C from 1 September 2014 to the date when the "whitish millet" appeared.

RNA Isolation, Library Construction, and Data Preprocessing

We chose the plant total RNA isolation kit (polysaccharides and polyphenolics-rich) (Huayueyang, Beijing, China) to extract total RNA from the litchi samples. A total of 5 μg of total RNA with an appropriate amount of Oligo-dT25 beads (Invitrogen, Carlsbad, CA, United States) was used for RNA-Seq library construction. First, the enriched mRNA was fragmented into short fragments and reverse transcribed into cDNA by random primers. Next, second-strand cDNA was synthesized using DNA polymerase I, RNase H, and dNTPs. The cDNA fragments were then purified by Agencourt AMPure XP (Beckman Coulter, Pasadena, CA, United States). The purified cDNA fragments were end-repaired, modified with poly (A) tails, and ligated to Illumina sequencing adapters. At last, the size-selected fragments were amplified and purified.

Libraries of 87 litchi accessions and 24 'Nuomici' leaf samples were sequenced with Illumina HiSeq^{TM} 2500 (BGI Co. Ltd., Shenzhen, China). The full RNA-Seq data have been submitted to the Sequence Read Archive (SRA) of the NCBI under accession PRJNA766599 and PRJNA766549.[1] Adaptors, low-quality sequences, and rRNA sequences were removed from raw reads. Then, the cleaned reads were mapped to the litchi genome[2] by TopHat2 (version 2.0.3.12). The eXpress (v1.5.1[3]) was used to calculate the read counts and fragments per kilobase of transcript per million mapped reads (FPKM) of each gene.

Genotype Analysis

Genome Analysis ToolKit, 4.0.6.0 (GATK) was used for SNP-calling (Van der Auwera et al., 2013). SNPs with single sample sequencing depth ≥3, missing rate less than 0.2, and the minor allele frequency (MAF, the frequency of occurrence of uncommon alleles in the population) more than 5% were identified as the high-quality filtered SNPs. ANNOVAR (Wang et al., 2010) (version 77) was applied for SNP/InDel annotation. The sequence difference matrix between individuals was obtained from the population SNP data, and a phylogenetic tree was constructed by the neighbor-joining method in the software Treebest (version 1.9.2) and MEGA 6.0.[4] The admixture model-based software Admixture (version 1.3.0) (Alexander et al., 2009) was used to estimate population structure and generate the Q matrix. The software SPAGeDi (Hardy and Vekemans, 2002) (version 1.5[5]) and the method described by Loiselle et al. (1995) were used to obtain the kinship matrix (K matrix).

Association Analysis Between Flowering Phenotypic and Population Single-Nucleotide Polymorphism Data

The association analysis software TASSEL (Bradbury et al., 2007) (version 3.0) was applied for the association analysis of flowering phenotypic and population SNP data, and four models were constructed: the general linear model (GLM model); the mixed linear model with Q matrix as covariate (Q model); mixed linear model with K matrix as covariate (K model), and mixed linear model with Q matrix and K matrix as covariates (Q + K model). The value of column "Symbol A," "Symbol B" in **Supplementary Table 1**, and the natural logarithm of the value of column "Symbol C" in **Supplementary Table 1** were used for association analysis. Q–Q plot was used to compare the distribution of the observed (actual) *p*-value and expected (theoretical) *p*-value of the models. Then, the Bonferroni correction was applied to determine the threshold of significant level (0.05/markers number). The quantitative trait loci (QTLs) were defined as the most strongly associated SNP (lead SNP, *p*-value < 1.00E-05) loci, and the genes that contained those QTLs were defined as AGs.

Association Analysis Between Gene Expression and the Flowering Phenotypic Data

The relationship between gene expression level of the screened chilling accumulation related AGs (flowering, hormone synthesis, hormone signal induction-related AGs) and the flowering phenotypic data (Symbol C in **Supplementary Table 1**) was calculated by linear regression using R.[6] For each gene,

[1] https://www.ncbi.nlm.nih.gov/sra

[2] http://121.37.229.61:82/

[3] https://pachterlab.github.io/eXpress/index.html

[4] http://www.megasoftware.net/

[5] http://ebe.ulb.ac.be/ebe/SPAGeDi.html

[6] http://www.r-project.org/

FPKM value was regressed as the dependent variable and chilling accumulation as the independent variable, and R^2 and significance values were calculated. Differentially expressed genes (DEGs) in the controlled temperature treatment were determined by using the OmicShare tools (edgeR[7]), and significant DEGs were restricted with FDR $\leq$0.05 and the absolute value of fold-change $\geq$ 1.5. Pathway enrichment analysis of the DEGs were performed using the OmicShare tools.[8] The expression of finally selected CAGs (p-value < 1.00E-05) was presented as a heat map diagram by integrative toolkit TBtools (Chen et al., 2020) (v1.075).

Quantitative Real-Time PCR Analysis

Total RNA was extracted by RNAprep Pure Plant Kit (Polysaccharides & Polyphenolics-rich) (Tiangen Biotech, China). A total of 1 μg total RNA was used for the synthesis of first-strand cDNA by the TransScript One-Step gDNA removal and cDNA synthesis SuperMix Kit (Transgen Biotech, China). Primers F/R for qRT-PCR were designed by Primer 5.0 software and synthesized at Sangon Co. Ltd. (Shanghai). The litchi homolog β-*actin* (NCBI GenBank accession number: HQ588865.1) was used as the reference gene. All the primers are shown in **Supplementary Table 2**. qRT-PCR was performed on a CFX96 real-time PCR machine (Bio-Rad, United States) with run as follows: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s. A 2 × RealStar Green Power Mixture (GenStar BioSolutions, China) was used in the qRT-PCR reactions, and each analysis was performed in two technical replicates and three biological replicates. The transcript quantification of the genes was performed in relation to β-*actin* and calculated using $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Expressions of candidate genes were presented as a heat map diagram using pheatmap package (see text footnote 6).

RESULTS

Flowering Phenotype of the Litchi Accessions

Traits associated with flowering phenotype of the 87 accessions were presented by "Symbol A," "Symbol B," and "Symbol C" (**Supplementary Table 1**). The accessions were subdivided into eight subcategories based on the time of panicle primordium emergence. Category 1–2 including 10 accessions were defined as early-flowering group. Category 3 comprising 6 accessions was the early-medium flowering group, while categories 4–5, including 11 accessions, were the medium-flowering group. Category 6 comprising 17 accessions was the medium-late-flowering group. The last group, categories 7–8 including 43 accessions were defined as the late-flowering one ("Symbol A" in **Supplementary Table 1**). "Symbol B" indicated the days required for floral induction. The time required for floral induction was quite different. For example, the time for the early-flowering 'Sanyuehong' was 54 days, while that of 'Dongliuhao' was 169 days. "Symbol C" indicated the chilling accumulation for floral induction. The chilling accumulation of the accessions was multiple, that in the early-flowering 'Sanyuehong' was about 600 times lower than in 'Dongliuyihao'. In general, these symbols indicated a multiplicity of flowering time and chilling accumulation levels in the accessions.

Throughput and Quality of RNA-Seq Libraries of the Litchi Accessions

We constructed RNA-Seq libraries of the 87 litchi accessions and performed transcriptome sequencing. As shown in **Supplementary Table 3**, 1.97×10^8–4.15×10^9 raw data were generated, and 1.10×10^6–2.62×10^6 clean reads were obtained from the libraries, with clean read ratios of 64.8–97.2%. We then mapped the clean reads to the litchi genome data, and the alignment rates were 58.1–77.8%.

Cluster Analysis of Litchi Accessions Based on Population Single-Nucleotide Polymorphism Data

A total of 98,155 high-quality SNP sites were obtained (**Supplementary Table 4**), and the SNP data of the population were used to obtain a sequence difference matrix among individuals, and a phylogenetic tree based on the neighbor-joining method was constructed. As shown in **Figure 1**, the early-flowering accessions were clustered together, so as the medium- and late-flowering ones. The results suggested that the accessions with similar flowering time possessed similar genetic background. Interestingly, 'Khom' (Lc13) and 'Kaleka' (Lc20), introduced from Thailand, have early flowering phenotype in southern China. However, their genetic backgrounds are closer to the medium-late or late-flowering accessions.

Association Analysis Between Flowering Phenotypic Traits and Population Single-Nucleotide Polymorphism Data

The correlation trend lines of the actual p-values and theoretical p-values obtained by Q + K model based on time of panicle primordium emergence ("Symbol A") and the days required for floral induction ("Symbol B") were closer to the expected line (linear, $y = x$). However, there is almost no useful information in relation to flowering. Hence, the Q + K model with the natural logarithm of the value of chilling accumulation ("Symbol C") was chosen to perform the associated analysis, referred to as SNP loci analysis related to chilling accumulation (**Supplementary Figure 1**). As a result, we identified 1,411 lead SNPs corresponding to 1,115 AGs (**Supplementary Table 5**).

Screening of Candidate Associated Genes in Relation to Chilling Accumulation for Floral Induction

From the 1,115 AGs, we detected 42 SNPs corresponding with 31 AGs in relation to flowering regulation (**Figure 2A** and **Supplementary Table 6**). A total of 16 genes were found to be related to photoperiod, circadian clock, and

[7]https://www.omicshare.com/tools/Home/Soft/diffanalysis
[8]http://www.omicshare.com/tools/Home/Soft/pathwaygsea

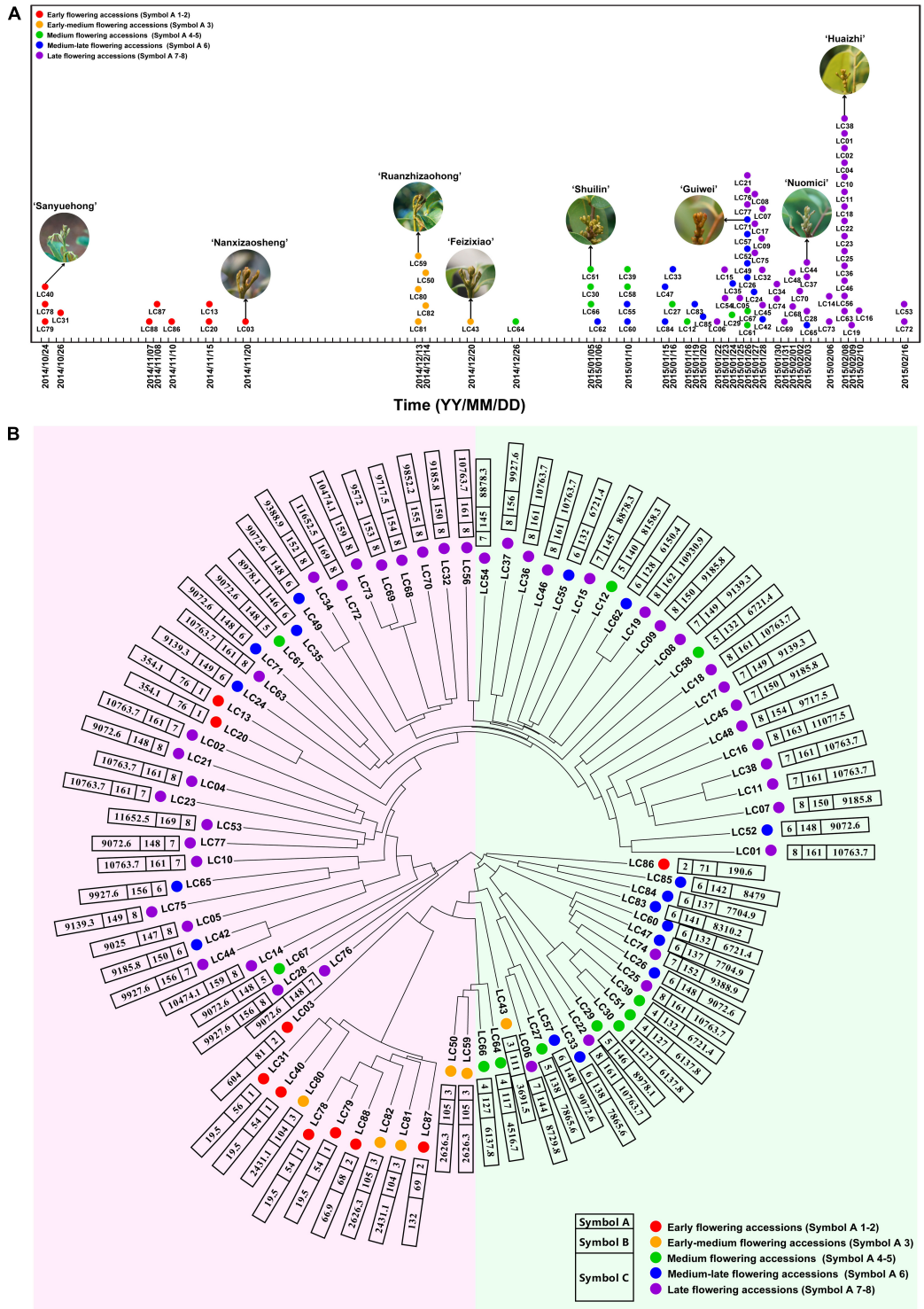


FIGURE 1 | Schematic diagram of flowering time and phylogenetic tree of litchi accessions based on SNP data. **(A)** Schematic diagram of flowering time of the 87 litchi accessions. **(B)** Phylogenetic tree of the accessions based on SNP data. Symbol A, flowering phenotype based on the time of panicle primordium ("whitish millet") emergence. Symbol 1–2, 3, 4–5, 6, 7–8 indicate early, early-medium, medium, medium-late, and late flowering accessions, respectively, and are signed by red, orange, green, blue, and purple dots. Symbol B, flowering phenotype based on days required for floral induction. Symbol C, flowering phenotype based on chilling accumulation. Detailed information of the accessions is listed in **Supplementary Table 1**.

light signaling pathway, including *CO* (CCG058352.1, CCG058353.1, CCG051351.1, and CCG067620.1), *GI* (CCG028841.1), *FKF1* (CCG038980.1), *ELF4* (CCG048202.1), *ZTL* (CCG005068.1 and CCG013580.2), *COP1* (CCG013698.1 and CCG031190.1), *NFYB3* (CCG029896.1), *PHYTOCHROMO C* (CCG045847.2 and CCG021941.1), *CKA2* (CCG033689.3), and *APRR9* (CCG028219.1). Five genes were related to the vernalization pathway, including *FRI3* (CCG021557.1), *FRI4* (CCG003706.2), *SUF4* (CCG019682.1), *VRN1* (CCG006818.1), and *EMF2* (CCG007444.2). Four aging pathway-related genes *SPL5* (CCG017709.1), *SPL12* (CCG053218.1), and *SPL14* (CCG047869.1 and CCG048984.2), and other six genes *ULT1* (CCG042033.1 and CCG060116.1), *UBC2* (CCG029021.2, CCG063563.1, and CCG066324.1), and *ESD4* (CCG037864.1) were identified.

Among the 1,115 AGs, 34 SNPs corresponding with 23 hormone synthesis-related AGs were identified, including 1 auxin (IAA), 1 cytokinin (CTK), 3 abscisic acid (ABA), 9 ethylene (ETH), 7 jasmonic acid (JA), and 2 salicylic acid (SA) synthesis related genes (**Figure 2B** and **Supplementary Table 6**).

A total of 27 plant hormone signal transduction AGs corresponding with 32 SNPs were identified, including 7 IAA, 1 CTK, 3 gibberellin (GA), 8 ABA, 5 ETH, 2 brassinosteroid (BR), and 1 JA signal transduction related genes (**Figure 2C** and **Supplementary Table 6**).

To further screen CAGs, we performed association analysis between gene expression levels of the AGs and the flowering phenotypic data. As a result, we identified 20 flowering-related CAGs out of those of the 31 AGs, 15 hormone biosynthesis-related CAGs out of the 23 AGs, and 13 hormone signal transduction-related CAGs out of the 27 AGs. Totally, 59% of the AGs were further screened (**Figure 2** and **Supplementary Table 7**).

Digital Transcriptomic Analysis of the Controlled Temperature-Treated Trees

To verify whether the CAGs are involved in flowering, we performed controlled temperature treatment with the 'Nuomici' litchi trees, a widely cultivated late-flowering variety. The result showed that the trees treated with low temperature (L) flowered 100%, while those treated with relatively high temperature (H) did not flower. We then performed RNA-Seq of the 24 leaf samples from the flowering and non-flowering treated trees. As shown in **Supplementary Table 8**, the clean reads of the samples were 3.69×10^6–9.75×10^6, and the alignment rates were all above 82%. We conducted a pairwise comparison of low- and high-temperature treatment at the same time point to identify DEGs using edgeR, and 17,727 DEGs were obtained, of which 8,782 DEGs were from the 3-day time point, 7,927 DEGs were from the 30-day time point, 9,036 DEGs were from the 60-day time point, and 4,180 DEGs were from the 75-day time point (**Supplementary Table 9**). The KEGG pathway enrichment analysis of the DEGs showed that the plant–pathogen interaction (ko04626), plant hormone signal transduction (ko04075), starch and sucrose metabolism (ko00500), RNA transport (ko03013), spliceosome (ko03040), RNA degradation (ko03018), carbon metabolism (ko01200), ribosome biogenesis in eukaryotes (ko03008), amino sugar and nucleotide sugar metabolism (ko00520), and protein processing in endoplasmic reticulum (ko04141) pathways were significantly enriched (**Supplementary Table 9**).

Expression Profiles of the Candidate Associated Genes in 87 Accessions and in Trees With Temperature-Controlled Treatment

An overlap of 797 genes was found from the CAGs of the transcriptome-based GWAS and DEGs in the controlled temperature experiment, during which 15 were flowering-related, 13 were hormone biosynthesis-related, and 11 were hormone signal transduction-related. Heat map diagram of the 15 flowering-related CAGs indicated *CO1* (CCG058352.1), *FKF1* (CCG038980.1), *PHYC* (CCG045847.2), *COP1* (CCG031190.1), *NFYB3* (CCG029896.1), *CKA2* (CCG033689.3), *APRR9* (CCG028219.1), *VRN1* (CCG006818.1), *EMF2* (CCG007444.2), *SPL12* (CCG053218.1), *UBC2* (CCG029021.2), *UBC2* (CCG063563.1), and *ULT1* (CCG042033.1) showed high expression levels in most of the late-flowering accessions, but low levels in most of the early-flowering and early-medium flowering accessions. *ZTL* (CCG013580.2), *SPL14* (CCG047869.1), and *UBC2* (CCG029021.2) showed low expression levels in most of the late-flowering accessions, but high expression levels in most of the early- and early-medium flowering accessions (**Figure 3A**). Interestingly, during the floral induction period (L3D, L30D, and L60D), expression patterns of the flowering-related CAGs in the flowering trees are contrary to those in the non-flowering ones (**Figure 3B**).

For the hormone synthesis-related CAGs, as shown in **Figure 4A**, the expression levels of the 13 genes in the early- and the early-medium flowering accessions are different from those in the other accessions. The CAGs related to the synthesis of IAA (CCG015050.1), ABA (CCG048171.1, CCG008586.1, and CCG055456.1), ETH (CCG047321.1 and CCG062263.2), and JA (CCG061887.1, CCG030510.1, CCG009077.1, CCG042227.1, and CCG032315.2) showed low-expression levels in most of the early- and the early-medium flowering accessions, with the exception of the 2 CAGs related to synthesis of CTK (CCG030382.1) and ETH (CCG045946.1). In the controlled temperature treatment, similar to the flowering-related CAGs, the expression pattern of the hormone synthesis-related CAGs in the low temperature-treated trees are contrary to those in the high temperature treated ones during the floral induction period (**Figure 4B**).

For the hormone signal transduction-related CAGs, the expression levels of the 11 CAGs in the early- and early-medium flowering accessions are quite different from the other accessions. The CAGs related to hormone signal transduction of IAA (CCG024350.1), ABA (CCG059706.1), ETH (CCG026609.1 and CCG047016.1), and BR (CCG034978.1) showed high expression levels in most of the early- and early-medium flowering accessions, whereas those related to hormone signal transduction of IAA (CCG067662.2, CCG008564.1, CCG032238.1, and

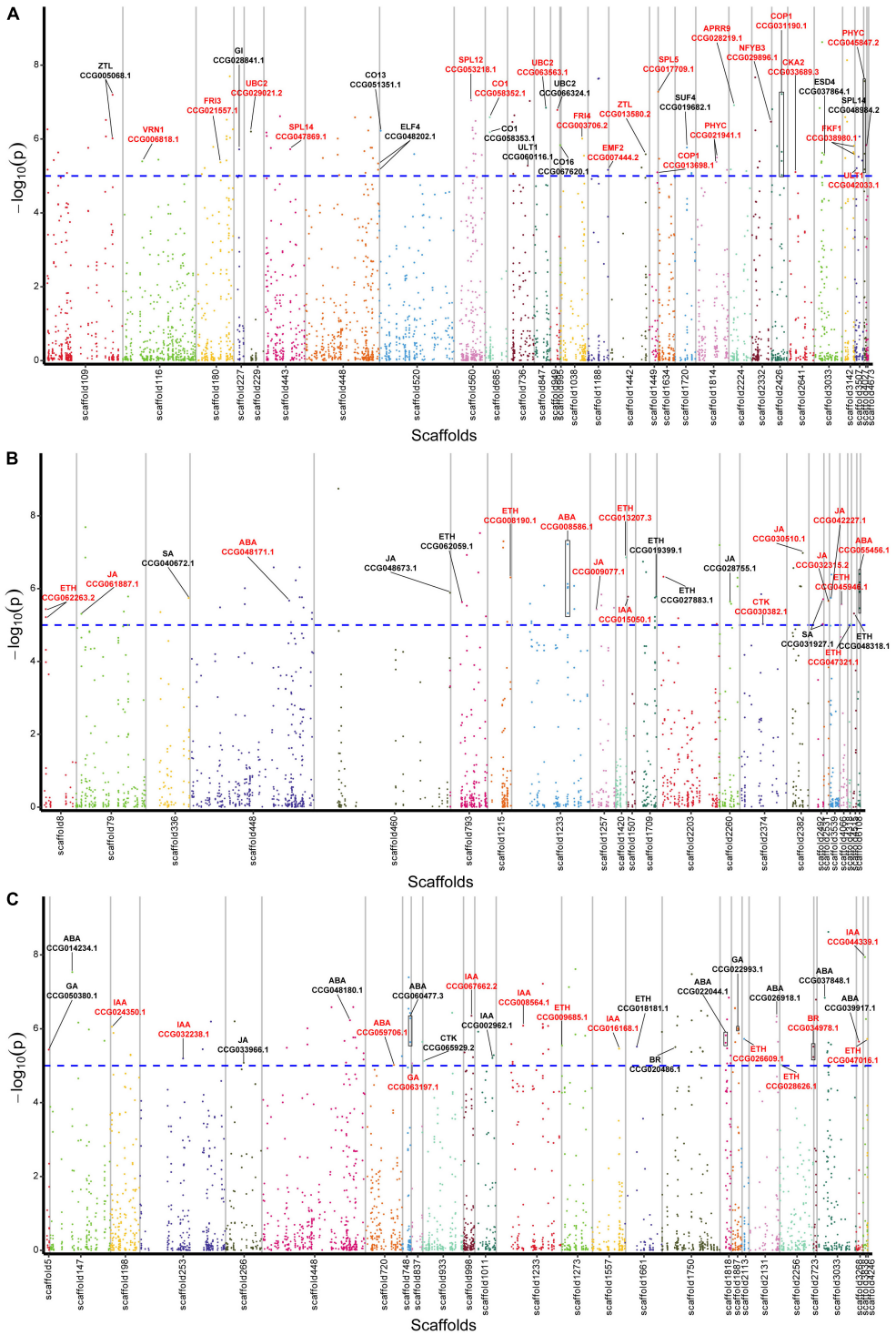


FIGURE 2 | Manhattan plot displaying the location of lead SNPs corresponding associated genes. **(A)** Flowering-related AGs. **(B)** Plant hormone biosynthesis-related AGs. **(C)** Plant hormone signal transduction-related AGs. The AGs are indicated with black lines, and the CAGs are marked by red color. Horizontal dashed line indicates significance threshold of $\log_{10}$ (p-values) = 5.

CCG044339.1) and ETH (CCG009685.1 and CCG028626.1) showed low expression levels (**Figure 5A**). In the controlled temperature treatment, the CAGs in the low temperature-treated trees are quite different from those in the high temperature treated ones. Most of the CAGs showed opposite expression levels between the low and the high temperature-treated trees.

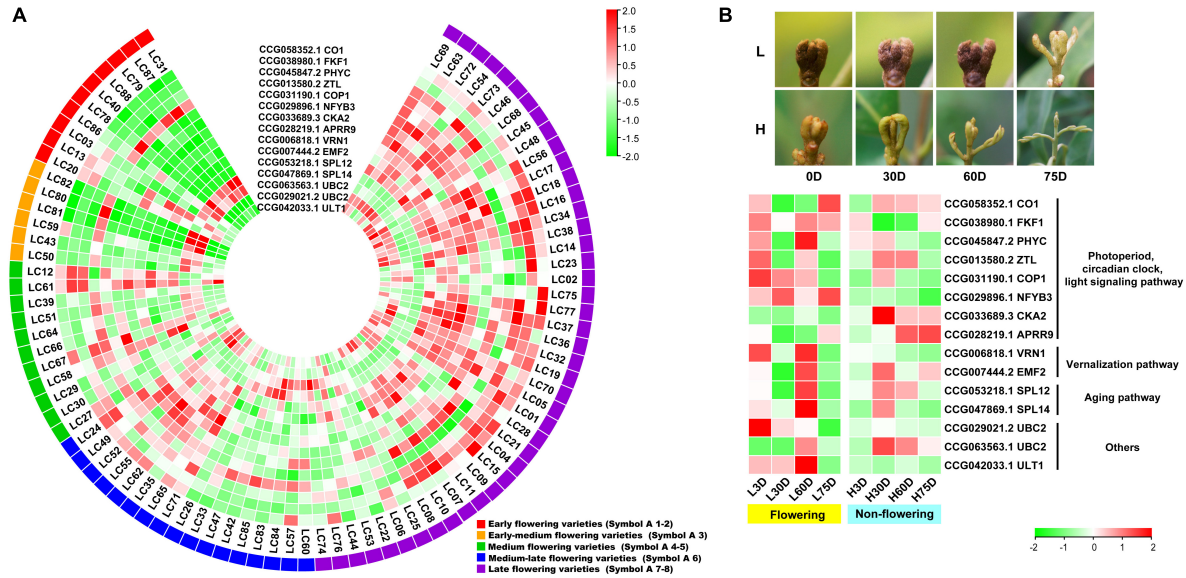


FIGURE 3 | Heat map diagram of the expression of the flowering-related CAGs. **(A)** Heat map of the expression levels of the flowering-related CAGs in the 87 litchi accessions. Symbol 1–2, 3, 4–5, 6, 7–8 indicate the early (red), the early-medium (orange), the medium (green), the medium-late (blue), and the late flowering accessions (purple), respectively. **(B)** Morphology of the buds and heat map diagram of the expression levels of the CAGs of the 'Nuomici' litchi trees under high temperature (25/20°C, day/night temperature, 12 h day and 12 h night) and low temperature (15/8°C, day/night temperature, 12 h day and 12 h night) conditions. L3D, L30D, L60D, and L75D indicate 3, 30, 60, and 75 days of the low-temperature treatment, respectively. H3D, H30D, H60D, and H75D indicate 3, 30, 60, and 75 days of high-temperature treatment, respectively. FPKM values were normalized to *Z*-score.

Meanwhile, except for genes related to the signal transduction of IAA (CCG032238.1 and CCG044339.1), most of the CAGs showed high expression levels during the floral induction period and decreased expression levels at the floral initiation stage. Also, they showed relatively low expression levels in the non-flowering trees (**Figure 5B**).

To confirm the accuracy of the transcriptome analysis results, we randomly selected 10 genes for qRT-PCR confirmation, and the results suggesting the transcriptome analysis by RNA-Seq is reliable (**Figure 6**).

Expression of the Candidate Associated Genes Under Gradient Chilling Conditions

To investigate the expression patterns of the CAGs under gradient chilling conditions, the medium-late-flowering accession, 'Guiwei' litchi trees were grown in a growth chamber at 10°C for 5 (Treatment 1), 7 (Treatment 2), and 9 (Treatment 3) weeks. A gradient increase in the flowering rate of the treatments was found. Trees of Treatment 1 did not flower, while those of Treatment 2 and 3 flowered with a percentage of flowering terminal shoots of 17.60 and 75.24%, respectively (**Figure 7A**).

Expressions of *CO1* and *ZTL* showed overall decreasing trends, whereas those of the *SPL12*, *UBC2*, *VRN1*, *SPL14*, and *ULT1* showed overall rising trends. *FKF1*, *PHYC*, and *APRR9* remained at low levels before S4, and increased sharply at S5, a time point for enough chilling accumulation for floral induction. In general, except for *CO1* and *ZTL*, expression levels of the flowering-related CAGs increased sharply at the time point when the trees obtained sufficient chilling accumulation for floral induction (**Figure 7B**).

Hypothesis of the Chilling Accumulation Related Genes in the Regulation of Flowering Time in Litchi

Based on the above-presented results, a potential mechanism of flowering time underlying chilling accumulation in litchi was proposed and presented in **Figure 8**. In leaves, under winter chilling or controlled chilling conditions, flowering-related genes such as the *FKF1* and *PHYC* promote *CO* accumulation, and the *APRR9*, *COP1*, *ZTL*, and *NFYB3* have a direct or indirect interaction with *CO*. CO activates the transcription of *FT*. In the phloem system, the FT protein acts as florigen, migrating from leaves to the SAMs to induce floral transition. Other flowering-related genes such as *SPL12*, *SPL14*, *ULT1*, and *UBC2* directly or indirectly promote *FT* expression. *VRN1* and *EMF2* may affect the expression of the *FLC* whose protein acts as a repressor to inhibit the expression of *FT*. Also, chilling may change the hormonal regulation system by affecting the expression of the plant hormone biosynthesis or signal transduction-related genes. The changes in the hormonal regulation system may on one hand produce hormone signals which can move from leaves to the apical meristems through vascular systems to promote floral transition, and on the other hand may promote the *FT* expression by several pathways. In the shoot apex, the migrated FT proteins and the hormone signals may synergistically promote floral transition.

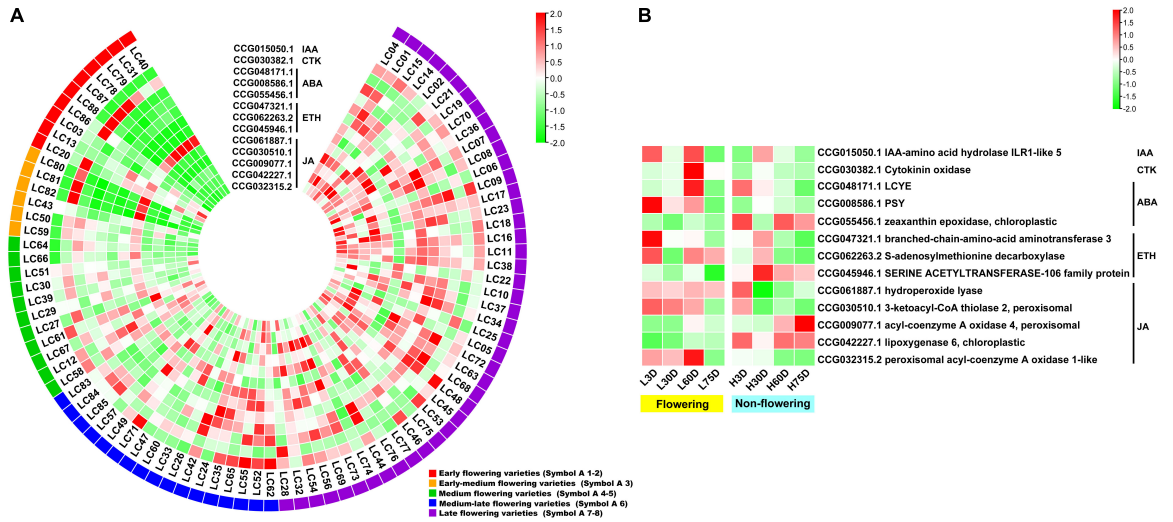


FIGURE 4 | Heat map diagram of the expression of hormone biosynthesis-related CAGs. **(A)** Heat map of the expression levels of the hormone synthesis-related CAGs in the 87 litchi accessions. **(B)** Heat map diagram of the expression levels of flowering-related CAGs under low and high temperature conditions. L3D, L30D, L60D, and L75D indicate 3, 30, 60, and 75 days of the low-temperature treatment, respectively. H3D, H30D, H60D, and H75D indicate 3, 30, 60, and 75 days of the high-temperature treatment, respectively. FPKM values were normalized to *Z*-score.

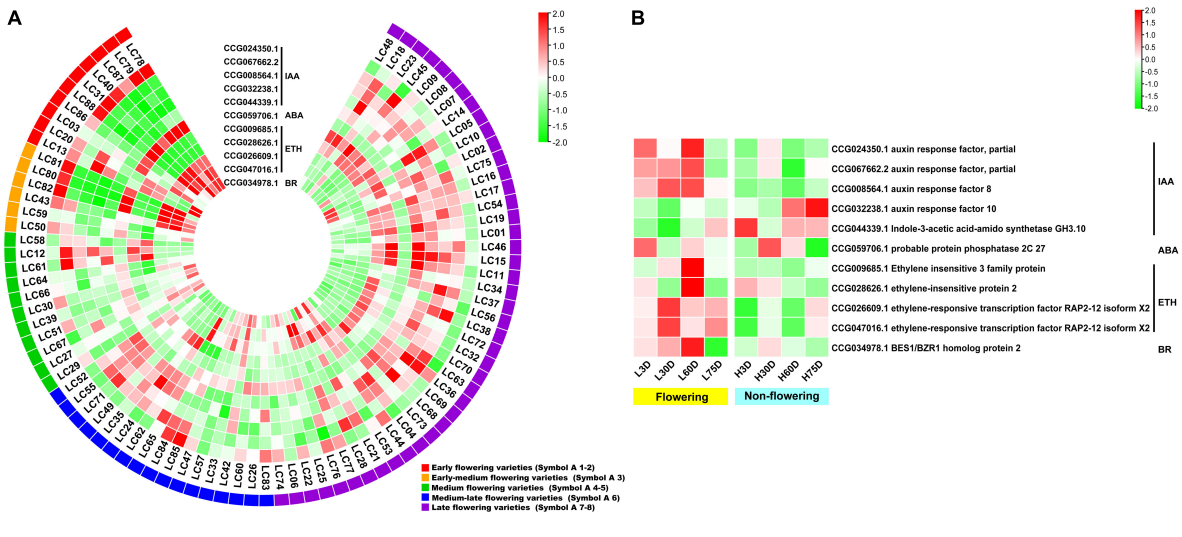


FIGURE 5 | Heat map diagram of the expression of the hormone signal induction-related CAGs. **(A)** Heat map of the expression levels of the hormone signal induction-related CAGs in the 87 litchi accessions. **(B)** Heat map diagram of the expression levels of flowering-related CAGs under low and high temperature conditions. L3D, L30D, L60D, and L75D indicate 3, 30, 60, and 75 days of the low-temperature treatment, respectively. H3D, H30D, H60D, and H75D indicate 3, 30, 60, and 75 days of the high-temperature treatment, respectively. FPKM values were normalized to *Z*-score.

DISCUSSION

Flowering is a transition from vegetative growth to reproductive growth. The complexity of this transition is created by an intricate network of signaling pathways which comprise diverse genes. Identification of these genes is important for crop production. In the past few years, genome-wide association studies (GWAS) have been widely used in rice (Huang et al., 2012), soybean (Zhou et al., 2015), maize (Tian et al., 2011), and melon (Pereira et al., 2018) for the identification of important trait-controlled genes. It has also been used to identify flowering-related genes in Arabidopsis (Aranzana et al., 2005), rice (Liu et al., 2021), maize (Maldonado et al., 2019), chrysanthemum (Chong et al., 2019), and soybean (Kim et al., 2020). Right now, no publication has reported any flowering time-related genes identified by GWAS in litchi. GWAS is useful for detecting potential targeted loci and candidate genes responsible for variations in complex traits (Bac-Molenaar et al., 2015; Sun et al., 2017). Transcriptome-based GWAS is a means of reducing sequence complexity from that represented by the entire genome

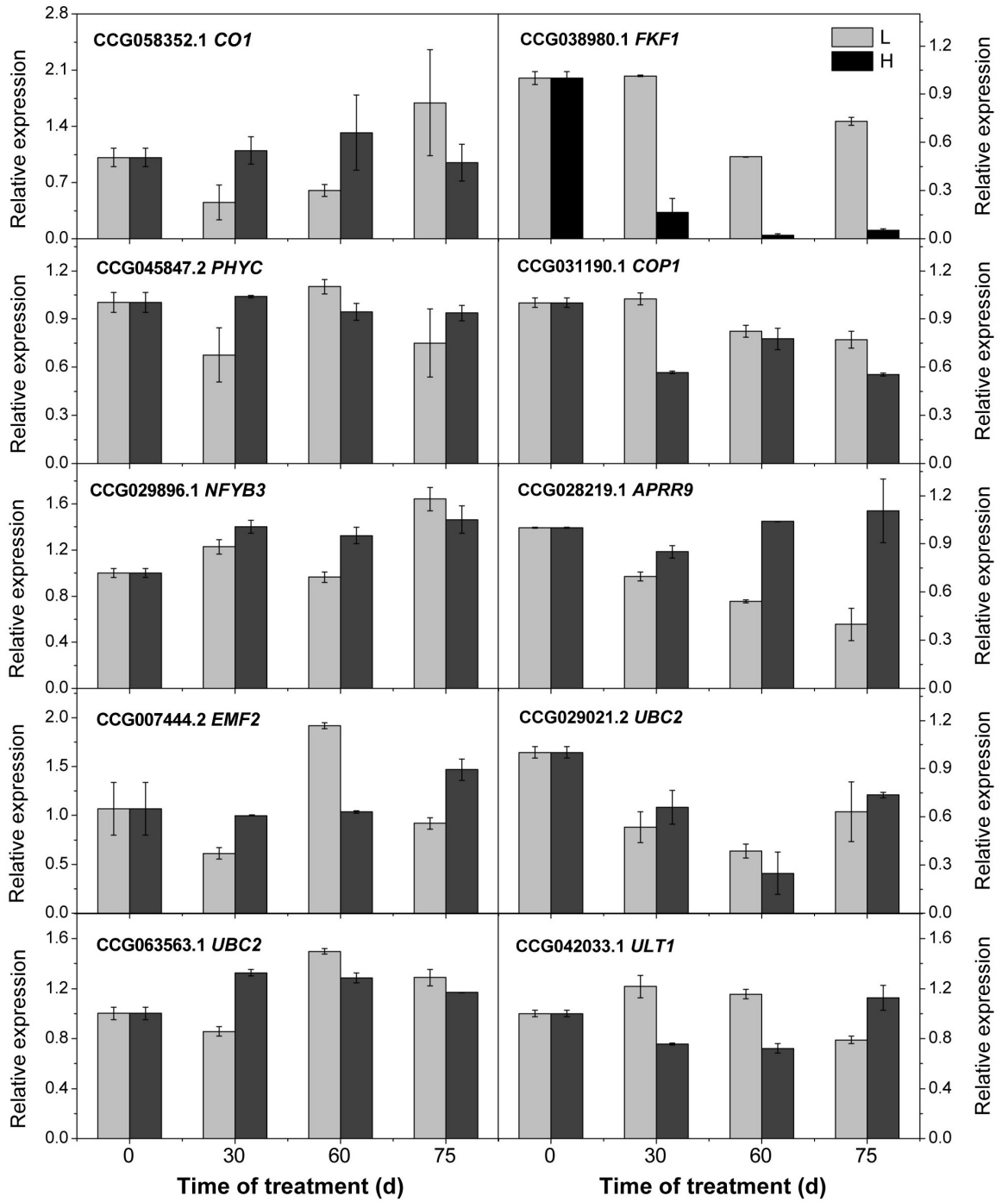


FIGURE 6 | Relative expression of the candidate genes revealed by quantitative real-time PCR. L and H indicate the low-temperature and high-temperature treatment, respectively. The data are the means of three biological replicates, and the bars represent SEs. The litchi homolog β-*actin* was used as a reference gene.

(Harper et al., 2012; Galpaz et al., 2018). With the advantages of simplifying analysis and reducing cost, transcriptome-based GWAS can also analyze the variation in different individuals (Zhu et al., 2018). Hence, in the present study, we used a transcriptome-based GWAS to identify flowering time-related genes in litchi.

Litchi is a woody fruit tree with a long-life cycle as well as a long juvenile phase. Time and cost are involved to construct hybrid populations for association analysis based on flowering-related phenotype. In the present study, we used a natural group of the collected accessions in our germplasm resource garden. The accessions were from China, Thailand,

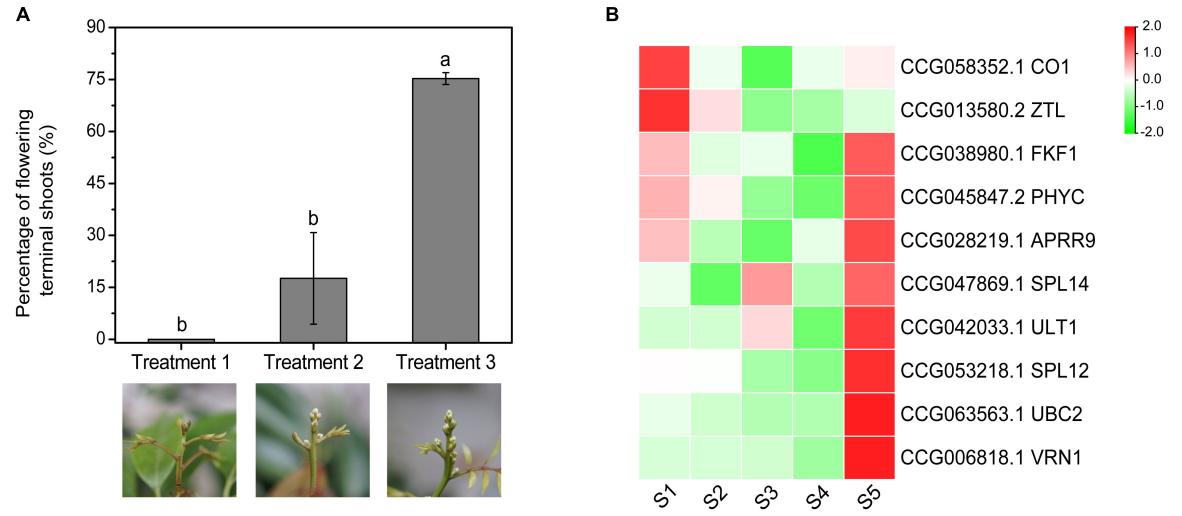


FIGURE 7 | Effects of gradient chilling treatments on flowering **(A)** and CAG expression **(B)** of litchi trees. 'Guiwei' potted trees were grown in a growth chamber at 10°C for floral induction. The trees were transferred to a growth chamber at 22°C after 5, 7, and 9 weeks, respectively. Different letters indicate significant differences at $p < 0.05$ among the treatments according to Tukey's multiple range test. S1, S2, S3, S4, and S5 indicate the time points of 0, 21, 35, 49, and 63 days of chilling.

South Africa, Australia, and Bangladesh with diverse flowering time. Also, they were planted at the same time in our germplasm resource garden to provide a similar growing environment and comparable flowering phenotypic data of the accessions. We constructed a phylogenetic tree, showing that most of the litchi accessions were clustered in accordance with the flowering time, suggesting that individuals with similar flowering time may have a similar genetic background. Interestingly, we found that the background of two early-flowering accessions, 'Khom' (Lc13) and 'Kaleka' (Lc20) are closer to those medium-late- or late-flowering accessions. Although they are introduced from Thailand where the ecological condition is different, they belong to the *L. chinensis*. Some mutations may happen so that they can flower under a warmer climate condition. Further study should be carried out to investigate the flowering mechanism of these interesting accessions.

For quantification of the specific floral induction time period into quantitative traits for association analysis, we used three kinds of flowering phenotypic traits. We finally identified lead SNPs using the data based on the natural logarithm of the value of degree hours (Symbol C) rather than those just based on time (Symbol A and B). Previous studies have shown that low temperature is required for litchi floral induction (Menzel and Simpson, 1988; Chen and Huang, 2005; Zhou et al., 2014). Chen et al. (2016) established a temperature model for litchi flowering and provided a base temperature of 19.32 for the sum of chilling accumulation. In our study, we used 20°C as the base temperature. We proposed that the base temperature should cover all the accessions including the early flowering accessions so as to get comparable phenotypic data. According to our previous study in the controlled temperature treatments, the early flowering accession 'Sanyuehong' could stably undergo floral induction in growth chambers at no more than 20°C (Liu, 2017). Hence, in the present study, we selected 20°C as the base temperature. Actually, the result showed that the chilling accumulation data were more workable for association analysis.

It is well studied that *FT* or its homolog encodes a long-seeking florigen protein that migrates from the leaves to the apical meristem to promote floral initiation (Corbesier et al., 2007; Yang et al., 2007). In the apical meristem, it interacts with the FLOWERING LOCUS D (FD) to promote floral transition and to initiate floral development through transcriptional activation of a floral meristem identity gene, *APETALA1* (*AP1*) (Abe et al., 2005). Ding et al. (2015) showed that *LcFT* expressed in leaves plays a pivotal role in litchi floral induction. As the leaves are part of the chilling sensing organs, and to produce signal for floral transition, in the present work, we focused on the chilling effects on the litchi leaves in which they produce flowering signals. We collected the leaf samples for the association and the transcriptomic analysis, and recorded the emergence time point of panicle primordia, which were visible as "whitish millets," symbolizations of the accomplishment for floral induction. Interestingly, our research on temperature-controlled treatment indicated that *LcFT1* (CCG039962.1) and *LcFT2* (CCG035605.1) in leaves of the low temperature-treated trees were also activated by the chilling treatment, whereas no expression levels of the genes were detected in the non-flowering trees (**Supplementary Table 9**), suggesting that *LcFT1/2* plays important roles in flowering. It may be that the sequence of *FT* is conserved in the litchi accessions and *FT* functions as a downstream integrated gene in flowering regulation, and we did not screen the *LcFT1/2* in the transcriptome-based GWAS. Huang and Chen (2005) pointed out that apical buds were in the dormant stage, and floral initiation occurs from this stage onward when cell division becomes active. There are three types of dormancy: endodormancy, paradormancy, and ecodormancy (Lang et al., 1987). In the temperate fruit tree almond, flowering is a complex process involving chilling and

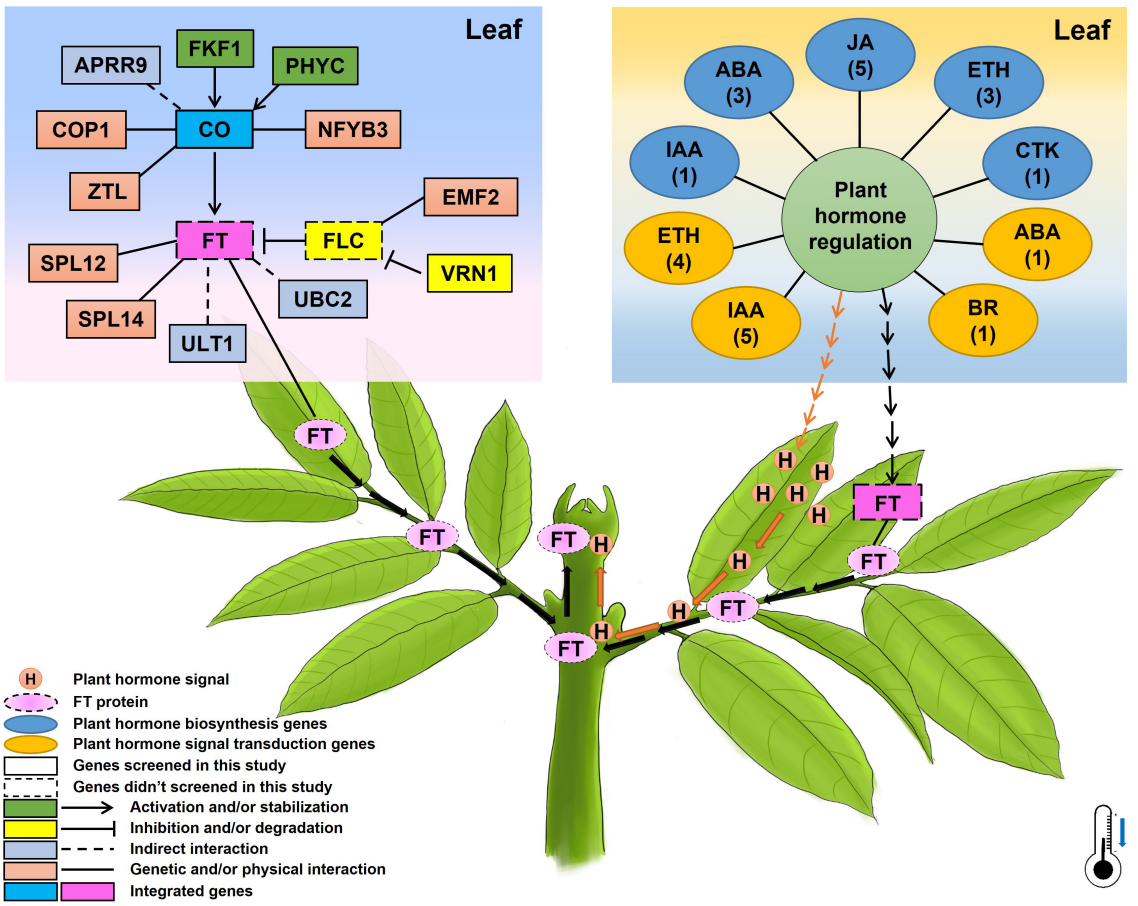


FIGURE 8 | A proposed diagram showing the possible genetic pathways regulating flowering time/floral transition underlying chilling accumulation in litchi. The numbers below the names of plant hormones in the oval boxes represent the members of the identified CAGs, and the thermometer with a blue arrow indicates the low temperature required for litchi flowering in winter.

heat requirements. The chilling requirements mainly affects endodormancy periods and are for breaking dormancy, and the heat requirements affect ecodormancy and are for flowering and productivity (Prudencio et al., 2018). Unlike deciduous temperate almond whose dormancy mechanism is well studied (Prudencio et al., 2019, 2021), there is less information about the evergreen litchi. Hieke et al. (2002) pointed out that the dormancy in litchi buds is not paradormancy or ecodormancy. It may be the endodormancy, and it is not induced or released by environmental cues but a purely endogenous process (Zhang et al., 2016). It might be that the chilling accumulation mainly induced the floral signals in leaves. However, how these signals affect the apical buds and promote the break out of the "whitish millets" still needs further investigation.

The GWAS is powerful for identifying candidate genes corresponding to natural variation of traits in field crops as well as in horticultural plants (Su et al., 2019; Alseekh et al., 2021; Berhe et al., 2021). A larger population size has been used for association analysis in field crops, such as rice (Huang et al., 2012), soybean (Zhou et al., 2015), and maize (Maldonado et al., 2019). GWAS has also been identified in trait-related candidate genes or SNPs in horticultural plants using populations less than 100, such as rose (Schulz et al., 2016), chestnut (Kang et al., 2019), chrysanthemums (Su et al., 2019), and bigleaf hydrangea (Wu and Alexander, 2020). In this research, we used 87 litchi accessions for the transcriptome-based GWAS and obtained 1,411 lead SNPs and 1,115 AGs. Then, we performed association analysis between gene expression and the flowering phenotypic data. Then a temperature-controlled experiment was carried out and 15 flowering-related CAGs, 13 hormone synthesis-related CAGs, and 11 hormone signal transduction-related CAGs were finally obtained. Interestingly, the expression levels of the 48 CAGs in the early and early-medium flowering accessions were quite different from the late flowering accessions (**Figures 3A**, **4A**, **5A**), and the possible reason might be laid on the different genetic back grounds as shown in the phylogenetic tree based on population SNP data (**Figure 1**). In this study, the lead SNPs were located on the CDS or 3′- or 5′-untranslated regions of the corresponding CAGs. The sequential difference of the CAGs between the early/early-medium flowering accessions and the late flowering accessions might affect their expression. Interestingly, most of the CAGs were highly expressed in the flowering trees and lowly

expressed in the non-flowering trees just at the same time points (**Figures 3B**, **4B**, **5B**), suggesting that expression patterns of these CAGs in the flowering trees were quite different from those in the non-flowering ones. The results further confirm that the CAGs are related to flowering time/chilling accumulation.

Plant hormones are signal molecules produced in plant cells in response to environment and are expressed at extremely low concentrations, but regulate a wide range of processes, including floral transition (Davis, 2009). We have identified IAA, ABA, ETH, CTK, BR, and JA synthesis-related or signal transduction-related CAGs. Interestingly, the DEGs in the controlled-temperature treatment were enriched in the plant hormone signal transduction pathway. Gimode et al. (2020) identified *PP2C* as a candidate flowering time gene in watermelon using QTL-Seq method. Phosphatases (DBP), which exhibit protein phosphatase 2C activities, are important regulators that are involved in both the transcriptional and posttranslational regulations. It can suppress *FLC* expression and enhance the expression levels of CO and *FT* in Arabidopsis (Zhai et al., 2016). We also identified an ABA signal transduction related gene *PP2C* (CCG059706.1). Brassinosteroids (BRs) are plant steroid hormones and play important roles in plant growth and development. The BR signal transduction components BES1/BZR1 can interact with two proteins, ELF6 (early flowering 6) and its homolog REF6 (relative of early flowering 6) to regulate flowering time in Arabidopsis (Yu et al., 2008). In accordance with the Arabidopsis, litchi flowering was proven to be regulated by ABA (Cui et al., 2013) and BR (Hu et al., 2018). The identified hormone-related CAGs may provide candidate genes for genetic regulation of flowering underlying hormonal control.

We also identified flowering-related CAGs whose homologs in the model plant Arabidopsis are proved to be involved in flowering, such as photoperiod or circadian clock regulation, and vernalization. Unlike the herbaceous Arabidopsis, litchi is an evergreen fruit tree whose flowering mechanism may be different. The known flowering regulation pathways in the model plant may be different from those in litchi. It is likely that the regarded photoperiod pathway related genes whose homologs in litchi may play different roles in flowering. In this study, those kinds of genes such as *CO1*, *FKF1*, *PHYC*, *ZTL*, and *COP1* may be involved in the regulation of flowering time underlying chilling accumulation in litchi. The identified candidate gene was *VRN1* (CCG006818.1), whose homolog plays an important role in vernalization and flowering in Arabidopsis (Levy et al., 2002). Its homolog in wheat, encoding a MADS-box transcription factor, could be significantly induced in leaves by vernalizations, and acts as the central regulator of wheat vernalization-induced flowering (Xie et al., 2021). *APRR9* (CCG028219.1) plays a general and an important role in the mechanisms underlying circadian rhythms in Arabidopsis. The APRR9-ox plants flowered much earlier than wild-type plants (Matsushika et al., 2002). On the whole, the expression of the CAGs showed significant difference between the early flowering and late flowering accessions. In addition, the expression profile of the CAGs showed different trends between flowering litchi trees and the non-flowering ones. Taken together, the screened CAGs may be related with flowering in litchi. Further functional studies should be carried out to investigate the role of the identified genes in litchi flowering underlying chilling accumulation.

As the late- or medium-late- flowering accessions need high chilling accumulation levels for floral induction, we supposed that flowering of these trees might have a dose-dependence on chilling accumulation, and the CAGs expression levels may change with the increase in chilling accumulation levels. We then performed an experiment in growth chambers. The medium-late-flowering accession, 'Guiwei' potted trees were grown under gradient chilling conditions. In accordance with our hypothesis, the flowering rates of the trees showed an increasing trend with the increase in chilling accumulation levels. Moreover, the expression of the flowering-related CAGs such as *FKF1*, *APRR9*, *PHYC*, *SPL12*, *SPL14*, *VRN1*, *ULT1*, and *UBC2* showed a sudden increase at the time point when the trees got sufficient chilling accumulation for flowering, suggesting these CAGs might need sufficient chilling accumulation for high expression.

CONCLUSION

We have performed transcriptome-based GWAS, obtained 1,411 lead SNPs and 1,115 AGs. To further screen CAGs, we performed association analysis between gene expression levels of the chilling accumulation related AGs and the flowering phenotypic data. We also performed a temperature-controlled experiment on litchi trees and explored the expression patterns of the screened CAGs in the flowering and non-flowering trees. Based on this conjoint analysis of RNA-Seq, we identified 15 flowering-related, 13 hormone biosynthesis-related, and 11 hormone signal transduction-related CAGs underlying chilling accumulation and flowering time in litchi. These genes may directly or indirectly affect the florigen or hormone signals produced in leaves that can migrate to the shoot apical apex to promote floral transition. Our present work for the first time provided CAGs by transcriptome-based GWAS in litchi. These genes may be applied for genetic regulation of flowering and for bypassing or partly bypassing chilling underlying climate change and global warming.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA766599; https://www.ncbi.nlm.nih.gov/, PRJNA766549.

AUTHOR CONTRIBUTIONS

BZ and SZ contributed to the design of the research and were contributors in writing the manuscript. XL performed sample collection, RNA isolation, library construction, association analysis, and writing the manuscript. PL carried out library construction and RNA-Seq data analysis. HL performed gene

expression analysis. LF performed RNA-Seq data analysis. HC participated in the design of the research. XP carried out data analysis. PXL performed the phenotyping. All authors read and approved the manuscript.

FUNDING

This research was funded by the National Natural Science Foundation of China (grant nos. 31772249 and 32072515), the National Litchi and Longan Research System (project no. CARS-32-07), and Hong Kong University Grants Committee (GRF-14111918 and AoE/M-403/16).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.819188/full#supplementary-material

REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., Ikeda, Y., et al. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309, 105–106. doi: 10.1126/science.1115983

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Alseekh, S., Kostova, D., Bulut, M., and Fernie, A. R. (2021). Genome-wide association studies: assessing trait characteristics in model and crop plants. *Cell. Mol. Life Sci.* 78, 5743–5754. doi: 10.1007/s00018-021-03868-w

Aranzana, M. J., Kim, S., Zhao, K. Y., Bakker, E., Horton, M., Jakob, K., et al. (2005). Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genet.* 1:e60. doi: 10.1371/journal.pgen.0010060

Bac-Molenaar, J. A., Vreugdenhil, D., Granier, C., and Keurentjes, J. J. B. (2015). Genome-wide association mapping of growth dynamics detects time-specific and general quantitative trait loci. *J. Exp. Bot.* 66, 5567–5580.

Berhe, M., Dossa, K., You, J., Mboup, P. A., Diallo, I. N., Diouf, D., et al. (2021). Genome-wide association study and its applications in the non-model crop *Sesamum indicum*. *BMC Plant Biol.* 21:283. doi: 10.1186/s12870-021-03046-x

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

Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, H. B., and Huang, H. B. (2005). Low temperature requirements for floral induction in lychee. *Acta Hortic.* 665, 195–202. doi: 10.17660/ActaHortic.2005.665.21

Chen, P., Roan, S., Lee, C., and Chen, I. (2016). Temperature model of litchi flowering-From induction to anthesis. *Sci. Hortic.* 205, 106–111. doi: 10.1016/j.scienta.2016.04.012

Chong, X. R., Su, J. S., Wang, F., Wang, H. B., Song, A. P., Guan, Z. Y., et al. (2019). Identification of favorable SNP alleles and candidate genes responsible for inflorescence-related traits via GWAS in chrysanthemum. *Plant Mol. Biol.* 99, 407–420. doi: 10.1007/s11103-019-00826-w

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., et al. (2007). FT protein movement contributes to long-distance signaling in floral induction of *Arabidopsis*. *Science* 316, 1030–1033. doi: 10.1126/science.1141752

Cui, Z., Zhou, B., Zhang, Z., and Hu, Z. (2013). Abscisic acid promotes flowering and enhances *LcAP1* expression in *Litchi chinensis* Sonn. *S. Afr. J. Bot.* 88, 76–79. doi: 10.1016/j.sajb.2013.05.008

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

Ding, F., Zhang, S., Chen, H., Su, Z., Zhang, R., Xiao, Q., et al. (2015). Promoter difference of *LcFT1* is a leading cause of natural variation of flowering timing in different litchi cultivars (*Litchi chinensis* Sonn.). *Plant Sci.* 241, 128–137. doi: 10.1016/j.plantsci.2015.10.004

Galpaz, N., Gonda, I., Shem-Tov, D., Barad, O., Tzuri, G., Lev, S., et al. (2018). Deciphering genetic factors that determine melon fruit-quality traits using RNA-Seq-based high-resolution QTL and eQTL mapping. *Plant J.* 94, 169–191. doi: 10.1111/tpj.13838

Gimode, W., Clevenger, J., and McGregor, C. (2020). Fine-mapping of a major quantitative trait locus Qdff3-1controlling flowering time in watermelon. *Mol. Breed.* 3:40. doi: 10.1007/s11032-019-1087-z

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., 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

Hieke, S., Menzel, C. M., and Lüdders, P. (2002). Shoot development, chlorophyll, gas exchange and carbohydrates in lychee seedlings (*Litchi chinensis*). *Tree Physiol.* 22, 947–953. doi: 10.1093/treephys/22.13.947

Hu, J., Kim, H., Chen, H., and Zhou, B. (2018). Litchi Flowering is Regulated by Expression of *Short Vegetative Phase* Genes. *J. Am. Soc. Hortic. Sci.* 143, 101–109. doi: 10.21273/JASHS04316-17

Huang, H. B., and Chen, H. B. (2005). A phase approach towards floral formation in lychee. *Acta Hort*. 665, 185–194.

Huang, X. H., Zhao, Y., Wei, X. H., Li, C. Y., Wang, A., Zhao, Q., et al. (2012). Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* 44, 32–53. doi: 10.1038/ng.1018

Kang, M. J., Shin, A. Y., Shin, Y., Lee, S. A., Lee, H. R., Kim, T. D., et al. (2019). Identification of transcriptome-wide, nut weight-associated SNPs in *Castanea crenata*. *Sci. Rep.* 9:13161. doi: 10.1038/s41598-019-49618-8

Kim, K. H., Kim, J. Y., Lim, W. J., Jeong, S., Lee, H. Y., Cho, Y., et al. (2020). Genome-wide association and epistatic interactions of flowering time in soybean cultivar. *PLoS One* 15:e0228114. doi: 10.1371/journal.pone.0228114

Lang, G. A., Early, J. D., Martin, G. C., and Darnell, R. L. (1987). Endo-, Para and ecodormancy: physiological terminology and classification for dormancy research. *HortScience* 22, 371–377.

Levy, Y. Y., Mesnage, S., Mylne, J. S., Gendall, A. R., and Dean, C. (2002). Multiple roles of Arabidopsis *VRN1* in vernalization and flowering time control. *Science* 297, 243–246. doi: 10.1126/science.1072147

Liu, C., Tu, Y., Liao, S. Y., Fu, X. K., Lian, X. M., He, Y. Q., et al. (2021). Genome-wide association study of flowering time reveals complex genetic heterogeneity and epistatic interactions in rice. *Gene* 770:145353. doi: 10.1016/j.gene.2020.145353

Liu, P. (2017). *Investigation on the Chilling Requirement of litchi Flowering and the Study of Gene Expression Patterns in Different Chilled-Type Cultivars*. M.Sc. thesis, Guangzhou: South China Agriculture University.

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

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

Maldonado, C., Mora, F., Bengosi Bertagna, F. A., Kuki, M. C., and Scapim, C. A. (2019). SNP- and haplotype-based GWAS of flowering-related traits in maize with network-assisted gene prioritization. *Agronomy Basel* 9:725. doi: 10.3390/agronomy9110725

Matsushika, A., Imamura, A., Yamashino, T., and Mizuno, T. (2002). Aberrant expression of the light-inducible and circadian-regulated *APRR9* gene belonging to the circadian-associated *APRR1/TOC1* quintet results in the phenotype of early flowering in *Arabidopsis thaliana*. *Plant Cell Physiol.* 43, 833–843. doi: 10.1093/pcp/pcf118

Menzel, C. M., and Simpson, D. R. (1988). Effect of temperature on growth and flowering of litchi (*Litchi chinensis* Sonn.) cultivars. *J. Hortic. Sci.* 63, 349–360. doi: 10.1080/14620316.1988.11515869

Pereira, L., Ruggieri, V., Perez, S., Alexiou, K. G., Fernandez, M., Jahrmann, T., et al. (2018). QTL mapping of melon fruit quality traits using a high-density GBS-based genetic map. *BMC Plant Biol.* 18:324. doi: 10.1186/s12870-018-1537-5

Prudencio, A. S., Díaz-Vivancos, P., Dicenta, F., Hernández, J. A., and Martínez-Gómez, P. (2019). Monitoring the transition from endodormancy to ecodormancy in almond through the analysis and expression of a specific Class III peroxidase gene. *Tree Genetics Genomes* 15:44. doi: 10.1007/s11295-019-1351-8

Prudencio, A. S., Hoeberichts, F., Dicenta, F., Martínez-Gómez, P., and Sánchez Pérez, R. (2021). Identification of early and late flowering time candidate genes in endodormant and ecodormant flower buds. *Tree Physiol.* 41, 589–605. doi: 10.1093/treephys/tpaa151

Prudencio, A. S., Martínez-Gómez, P., and Dicenta, F. (2018). Evaluation of breaking dormancy, flowering and productivity of extra-late and ultra-late flowering almond cultivars during cold and warm seasons in South-East of Spain. *Sci. Hortic.* 235, 39–46. doi: 10.1016/j.scienta.2018.02.073

Schulz, D. F., Schott, R. T., Voorrips, R. E., Smulders, M. J. M., Linde, M., and Debener, T. (2016). Genome-wide association analysis of the anthocyanin and carotenoid contents of rose petals. *Front. Plant Sci.* 7:1798. doi: 10.3389/fpls.2016.01798

Su, J. S., Zhang, F., Chong, X. R., Song, A. P., Guan, Z. Y., Fang, W. M., et al. (2019). Genome-wide association study identifies favorable SNP alleles and candidate genes for waterlogging tolerance in *chrysanthemums*. *Hortic. Res.* 6:21. doi: 10.1038/s41438-018-0101-7

Su, Z., Yang, S., Chen, H., and Shen, J. (2020). Analysis of the production situation for litchi in main planting areas of China in 2020. *J. S. Agri.* 51, 1598–1605. doi: 10.3969/j.issn.2095-1191.2020.07.012

Sun, Z. W., Wang, X. F., Liu, Z. W., Gu, Q. S., Zhang, Y., Li, Z. K., et al. (2017). Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. *Plant Biotechnol. J.* 15, 982–996. doi: 10.1111/pbi.12693

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, 113–159. doi: 10.1038/ng.746

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* 11:43. doi: 10.1002/0471250953.bi1110s43

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164. doi: 10.1093/nar/gkq603

Wu, X., and Alexander, L. W. (2020). Genome-wide association studies for inflorescence type and remontancy in *Hydrangea macrophylla*. *Hortic. Res.* 7:27. doi: 10.1038/s41438-020-0255-y

Xie, L., Zhang, Y., Wang, K., Luo, X. M., Xu, D. A., Tian, X. L., et al. (2021). *TaVrt2*, an *SVP-like* gene, cooperates with *TaVrn1* to regulate vernalization-induced flowering in wheat. *New Phytol.* 231, 834–848. doi: 10.1111/nph.16339

Yang, Y. J., Klejnot, J., Yu, X. H., Liu, X. M., and Lin, C. T. (2007). Florigen (II): it is a mobile protein. *J. Integr. Plant Biol.* 49, 1665–1669. doi: 10.1111/j.1744-7909.2007.00614.x

Yu, X., Li, L., Guo, M., Chory, J., and Yin, Y. (2008). Modulation of brassinosteroid-regulated gene expression by jumonji domain-containing proteins ELF6 and REF6 in *Arabidopsis*. *PNAS* 105, 7618–7623. doi: 10.1073/pnas.0802254105

Zhai, H., Ning, W., Wu, H., Zhang, X., Lu, S., and Xia, Z. (2016). DNA-binding protein phosphatase AtDBP1 acts as a promoter of flowering in *Arabidopsis*. *Planta* 243, 623–633. doi: 10.1007/s00425-015-2433-y

Zhang, H., Li, H., Lai, B., Xia, H., Wang, H., and Huang, X. (2016). Morphological characterization and gene expression profiling during bud development in a tropical perennial, *Litchi chinensis* Sonn. *Front. Plant Sci.* 7:1517. doi: 10.3389/fpls.2016.01517

Zhang, H., Shen, J., Wei, Y., and Chen, H. (2017). Transcriptome profiling of litchi leaves in response to low temperature reveals candidate regulatory genes and key metabolic events during floral induction. *BMC Genom.* 18:363. doi: 10.1186/s12864-017-3747-x

Zhou, B. Y., Chen, H. B., and Wu, C. B. (2014). An overview on natural triggers and stress signals in relation to flowering in *Litchi chinensis* and *Dimocarpus longan*. *Acta Hortic.* 1029, 137–144. doi: 10.17660/actahortic.2014.1029.15

Zhou, Z. K., Jiang, Y., Wang, Z., Gou, Z. H., Lyu, J., Li, W. Y., et al. (2015). Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat. Biotechnol.* 33, 125–408. doi: 10.1038/nbt0416-441c

Zhu, W., Xu, C., Zhang, J. G., He, H., Wu, K. H., Zhang, L., et al. (2018). Gene-based GWAS analysis for consecutive studies of GEFOS. *Osteoporosis Int.* 29, 2645–2658. doi: 10.1007/s00198-018-4654-y

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

HYPOTHESIS AND THEORY
published: 25 February 2022
doi: 10.3389/fpls.2022.804968

Comparative Transcriptome Analysis Reveals Differential Regulation of Flavonoids Biosynthesis Between Kernels of Two Pecan Cultivars

Chengcai Zhang, Huadong Ren, Xiaohua Yao*, Kailiang Wang and Jun Chang*

Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou, China

Flavonoids influence the flavor and nutritional value of pecan nuts. However, limited information is available regarding the molecular mechanisms underlying pecan flavonoid biosynthesis. Here, we used a high ("YLC28") and a low ("Oconee") flavonoid content cultivar as the research objects. The changes in flavonoid content and the gene transcription patterns during kernel development were identified. Different accumulation patterns of total flavonoids (TF) and condensed tannins (CT) were observed between the two cultivars. The contents of TF and CT in "YLC28" were 1.76- and 2.67-fold higher levels than that of "Oconee" on 150 days after full bloom of female flowers, respectively. In total, 30 RNA-Seq libraries were constructed and sequenced. The upregulated genes in "YLC28" were highly enriched in flavonoid-related pathways. Thirty-three structural genes were identified, and the expression of two *phenylalanine ammonia lyases*, one *chalcone synthase*, one *flavonoid 3',5'-hydroxylase*, and one *flavonol synthase* exhibited high correlation ($r \geq 0.7$, $p < 0.01$) with the condensed tannin content in "YLC28." A putative MYB transcription factor, CIL1093S0100, might act as a flavonoid biosynthesis repressor during kernel development. Altogether, these results will be useful for uncovering the molecular mechanisms of flavonoid biosynthesis and subsequently accelerating quality pecan breeding.

Keywords: flavonoid, condensed tannin, RNA-Seq, MYB, carya illinoinensis

OPEN ACCESS

Edited by:
Yuexue Liu,
Shenyang Agricultural University, China

Reviewed by:
Ze Peng,
South China Agricultural University, China
Dev Paudel,
University of Florida, United States

****Correspondence:***
Huadong Ren
renhd@163.com
Xiaohua Yao
yaoxh168@163.com

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

Received: *29 October 2021*
Accepted: *07 February 2022*
Published: *25 February 2022*

Citation:
Zhang C, Ren H, Yao X, Wang K and Chang J (2022) Comparative Transcriptome Analysis Reveals Differential Regulation of Flavonoids Biosynthesis Between Kernels of Two Pecan Cultivars. Front. Plant Sci. 13:804968. doi: 10.3389/fpls.2022.804968

INTRODUCTION

Pecan [*Carya illinoinensis* (Wangenh.) K. Koch] is native to North America, belongs to the Juglandaceae family and is an important nut crop throughout the world (Grauke et al., 2016). Its kernels are rich in unsaturated fatty acids, phenolics, flavonoids, condensed tannins, ellagic acid, and other bioactive components and are good for human health

Abbreviations: DEG, differentially expressed gene; PAL, phenylalanine ammonia lyase; CHS, chalcone synthase; CHI, chalcone isomerase; DFR, dihydroflavonol 4-reductase; TF, transcription factor; PA, proanthocyanidin; C4H, cinnamate 4-hydroxylases; TF, total flavonoid; CT, condensed tannin; FPKM, fragment per kilobase of transcript per million mapped reads; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; qRT-PCR, quantitative real-time PCR; SE, standard error; 4CL, 4-coumarate-CoA ligases; F3H, flavanone 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylases; LAR, leucoanthocyanidin reductases; ANS, anthocyanidin synthase; ANR, anthocyanidin reductase; FLS, flavonol synthase; EAR, ethylene-responsive element-binding factor-associated amphiphilic repression.

(Gong and Pegg, 2017; Zhang et al., 2019a). Among the different types of phytochemicals, phenolics are the most universal antioxidants in plants. Phenolics have the ability to neutralize harmful reactive oxygen species in cells and are excellent health-promoting components for humans (Caltagirone et al., 2000; Shay et al., 2015). Phenolics are key flavor and nutrition components associated with the bitterness, astringency, color, flavor, odor, and oxidative stability of foods (Pandey and Rizvi, 2009). Previous studies reported that pecan nuts contain high phenolic concentrations among the different phenolic richest foods (Kornsteiner et al., 2006; Pérez-Jiménez et al., 2010). Since Senter et al. (1980) first isolated 8 phenolic acids using gas-liquid chromatography-mass spectrometry, different classes of phenolic acids, ellagitannins, flavonoids, proanthocyanidins (condensed tannins) and anthocyanins have been successively identified in pecan nuts (Polles et al., 1981; Harnly et al., 2006; de la Rosa et al., 2011; Robbins et al., 2014; Gong and Pegg, 2017). The contents of phenolic components in pecan kernels exhibited intercultivar variation (Venkatachalam and Sathe, 2006; Villarreal-Lozoya et al., 2007; Zhang et al., 2019b; Bouali et al., 2020). For example, the tannin contents were 0.66 g/100 g in "Desirable" and 2.68 g/100 g in a Texas seedling; the content in these two genotypes differed by 4.06-fold (Venkatachalam and Sathe, 2006). The contents of total phenolics and condensed tannins showed at most 1.71- and 1.88-fold differences among the six pecan cultivars, respectively (Villarreal-Lozoya et al., 2007). Because phenolics are the major antioxidant phytochemicals in pecan (Villarreal-Lozoya et al., 2007; de la Rosa et al., 2011), the nuts from different varieties present different nutritive values and different implications for health promotion (Bouali et al., 2020). Therefore, the phenolic content and composition in kernels are key traits for pecan breeding. Verifying the genetic regulation of its metabolism will be helpful for understanding the quality formation mechanism and will promote genetic improvement in this plant.

Flavonoids are the main type of phenolics, and approximately 6000 compounds have been identified in plants. Several flavonoid compounds, such as condensed tannins, apigenin and quercetin, are beneficial for human health (Caltagirone et al., 2000). The biosynthesis mechanisms of flavonoids have been well characterized in model plants, which are synthesized by a series of enzymes, such as phenylalanine ammonia lyase (PAL), chalcone synthase (CHS), chalcone isomerase (CHI), and dihydroflavonol 4-reductase (DFR). In addition, the structural genes of these enzymes are transcriptionally regulated by several types of transcription factors, of which the most concerning is *R2R3-MYBs* (Ma and Constabel, 2019). TT2 (AtMYB123) is the first identified proanthocyanidin-related R2R3-MYB, which positively regulates the accumulation of proanthocyanidin in *Arabidopsis* seeds (Nesi et al., 2001). In contrast, AtMYB4 acts as a negative regulator of the phenylpropanoid pathway in *Arabidopsis* by negatively controlling the expression of *cinnamate 4-hydroxylases* (*C4H*) (Jin et al., 2000). To date, the functions of a large number of flavonoid-related MYBs have been identified in different plant species (Akagi et al., 2009; Ma and Constabel, 2019). However, knowledge about the structural genes and transcription factors associated with flavonoid biosynthesis in pecan remains limited (Zhang et al., 2019a).

RNA-Seq is a high-throughput sequencing method approach to gene expression profiling. To date, this technology has been widely utilized to investigate the genetic clues behind secondary metabolite biosynthesis (Liu et al., 2015; Wei et al., 2016; Li et al., 2018). Comparative transcriptome analysis is a powerful strategy to reveal the different regulatory mechanisms underlying flavonoid-content phenotypes in different varieties. Liu et al. (2015) compared the transcriptomes of white (normal) and purple (anthocyanin-rich) potato and identified a set of highly expressed anthocyanin-related genes in the purple genotype. Li et al. (2017) compared the difference in transcription patterns between red and green walnut and then identified structural genes and transcription factors associated with anthocyanin accumulation in red walnut. Similarly, genes involved in flavonoid biosynthesis and accumulation in *Camellia sinensis* (Wei et al., 2016), *Zanthoxylum bungeanum* (Sun et al., 2019), blueberry (Lin et al., 2018), and peanut were identified (Huang et al., 2019). Zhang et al. (2019a) determined the gene expression profiles in four developmental stages of pecan kernels, and 36 candidate unigenes associated with flavonoid synthesis were obtained. Three *Carya illinoinensis CHSs* (*CiCHSs*) were isolated, and two of them exhibited significant correlations with the total phenolic content of pecan (Zhang et al., 2019b). Seven *CiPALs* were identified in the pecan genome in which *CiPAL4* and *CiPAL5* might play crucial roles in phenolic biosynthesis in pecan (Zhang et al., 2022). Recently, a draft genome of pecan was generated, and a set of polyphenol-related candidate genes was reported (Huang Y. et al., 2019), and several of them show alternative splicing during pecan kernel development (Zhang et al., 2021). However, the functions and transcriptional regulation of these genes during flavonoid biosynthesis in pecan remain unclear.

In the present study, the different molecular regulation of flavonoid biosynthesis was explored in pecan kernels between a low ("Oconee") and a high flavonoid content cultivar ("YLC28"). The content variation of total flavonoids, condensed tannins, quercetin, and apigenin along with embryo development were determined. The gene expression profiles during kernel maturation of the two cultivars were identified by using RNA-Seq technology. Structural genes and transcription factors associated with flavonoid biosynthesis were obtained. The investigation of molecular mechanisms behind phytochemical accumulation based on comparative transcriptome was rarely performed in pecan. We believe that the findings will deepen our understanding of flavonoid accumulation in pecan kernels and provide ideas for the study of molecular mechanisms underlying other phytochemical biosynthesis in this plant. Subsequently, the present study will facilitate breeding research in pecan.

MATERIALS AND METHODS

Plant Material

The pecan kernels were sampled at approximately 100, 107, 114, 121, 128, 135, 142, and 150 days after the full bloom of

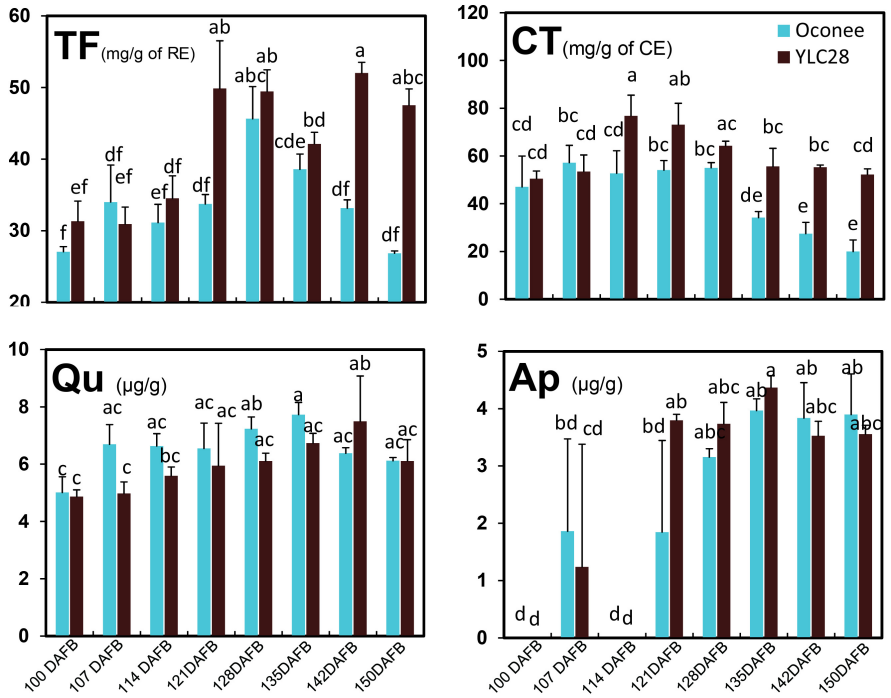


FIGURE 1 | The content changes of phenolic components during kernel development in the two cultivars. Total flavonoid (TF), condensed tannin (CT), quercetin (Qu), and apigenin (Ap) contents. The experiments were performed with three biological repetitions, and the error bars are presented as the mean ± standard error (SE).

female flowers (DAFB, 50% stigmas among 20 randomly selected clusters secreted stigmatic fluid) from "YLC28" (high flavonoid content, H) and "Oconee" (low flavonoid content, L) trees in Jiande (29° N, 119° W), China. The trees were 11-years old. The kernels were frozen in liquid nitrogen and stored at −80°C for further analysis.

Flavonoid Content Determination

The plant materials were dried in a freeze dryer (FreeZone, Labconco Corp., Kansas City, MO, United States) for 96 h. Then, the kernels were finely ground in a tissue blender (JYL-D022, Joyoung Co. Ltd., Shandong, China). The powder of each sample was used to detect the contents of major classes of flavonoids, such as total flavonoids, condensed tannins, quercetin, and apigenin.

The total flavonoids were extracted using an ethanol reflux method (Wang et al., 2017). 0.5 g powder of each sample was extracted with 30 mL of 80% ethanol two times. The filtrates were combined and evaporated to dryness. Then the extractives were dissolved by 30% ethanol and completed the total volume up to 15 mL. 1 mL of the extract was transferred to a 10 mL volumetric flask and incubated for 6 min after mixture with 0.5 mL 0.05 g/mL $NaNO_2$. Then, 0.3 mL 0.1 g/mL $Al(NO_3)_3$ was added, incubated for six min before mixture with 3 mL 0.04 g/mL NaOH, and added distilled water to make the total volume up to 10 mL. After incubating for 15 min, the absorbance was detected at 510 nm using a spectrophotometer. The content of TF was expressed as rutin equivalent (RE).

The CT content was detected using a vanillin assay (de la Rosa et al., 2011). 1.0 g powder of pecan kernels was placed in a 50 mL centrifuge tube. Afterward, 20 mL distilled water was added into the centrifuge tube. Then, the mixture of powder and distilled water was incubated in boiling water for 30 min. The supernatant was transferred into a new 50 mL volumetric flask. Then, the precipitate was re-extracted two times by 10 mL distilled water. The supernatants were combined, and added distilled water to make the total volume up to 50 mL. 10 mL extract was placed in a new 50 mL centrifuge tube, then freeze-dried and dissolved in 10 mL anhydrous methanol. After that, the extract was centrifuged, and the supernatant was used for the CT determination as described by de la Rosa et al. (2011). The content of CT was expressed as catechin equivalent (CE).

The standards of quercetin (Qu) and apigenin (Ap) were obtained from Sigma-Aldrich (Shanghai, China). The pecan kernel powder (0.5 g) was weighed, mixed with 1.5 mL of extracting solution and incubated at 80°C for 90 min, then extracted in an ultrasonic bath for 60 min. The mixtures were centrifuged (4°) at 10,000 r/min for 10 min and filtered through a 0.45 μm filter membrane. Finally, the extractives were dried under nitrogen flow, dissolved in 1000 μL methanol, and filtered (0.22 μm). HPLC was performed on a Rigol L3000 HPLC system (RIGOL Technology Co., Ltd., Beijing, China) with a Kromasil C18 column (250 mm × 4.6 mm, 5 μm particle size) at 35°C. The mobile phases were methanol/water/phosphoric acid (60/40/0.08). The flow rate was 0.8 mL/min. The detection wavelength was 360 nm. The contents of quercetin and apigenin were calculated using regression equations.

RNA Extraction, Library Construction, and Sequencing

To identify the molecular mechanism that regulates the different flavonoid contents between "YLC28" and "Oconee," RNA-Seq technology was employed. Based on the trends of TF and CT content variation, five different developmental stages (100 DAFB, 114 DAFB, 121 DAFB, 135 DAFB, and 150 DAFB) of kernels from the two cultivars were used as plant materials. The number of "YLC28" (high flavonoid content, H) samples ranged from H100 (100 DAFB) to H150 (150 DAFB). Similarly, the number of "Oconee" samples (low flavonoid content, L) ranged from L100 to L150. Total RNA was extracted using a TRIzol reagent kit (Invitrogen, Carlsbad, CA, United States) according to the manufacturer's protocol. RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States) and assessed using RNase-free agarose gel electrophoresis. Library construction was performed as described previously. Next, the 30 separately constructed libraries (five stages with three bioreplicates of two cultivars) were sequenced using Illumina HiSeq 2500 by Gene Denovo Biotechnology Co. (Guangzhou, China).

Data Processing

Data quality control, clean read assembly and gene abundance calculation were performed as described in a previous study (Shao et al., 2019). The raw reads were filtered by fastp (Chen et al., 2018) to obtain clean reads. Those reads containing adapters, unknown nucleotides higher than 10%, and low-quality bases (Q-value $\leq$ 20) higher than 50% were removed. An index of the reference genome was built (Huang Y. et al., 2019), and paired-end clean reads were mapped to the reference genome using HISAT2.2.4 with "-rna-strandness RF" and other parameters set as a default (Kim et al., 2015; Huang Y. et al., 2019). The mapped reads of each sample were assembled using StringTie v1.3.1 in a reference-based approach (Pertea et al., 2015). To identify novel genes, Cufflinks software (Trapnell et al., 2012) was employed and analyzed as described by Shao et al. (2019). For each transcription region, a fragment per kilobase of transcript per million mapped reads (FPKM) value was calculated to quantify its expression abundance and variations. In order to evaluate the replicates of the transcriptome samples, a principal component analysis (PCA) was performed using FPKMs of all genes by a SIMCA 14.1 software (V14.1, MKS Data Analytics Solutions, Umea, Sweden). The Gene Ontology (GO[1]) term annotation for each gene was performed by Blast2GO software[2] using default parameters, taking FDR $\leq$ 0.05 as a threshold. All the sequences were aligned to the KEGG database,[3] and PlantTFdb database[4] by BlastX (e value $<$ 1e-5).

A differential expression gene (DEG) analysis was performed by using DESeq2 software (Love et al., 2014). DEGs were selected with criteria of $|\log 2(\text{fold change})| \geq 1$, false discovery rate (FDR) $\leq$ 0.05. Then, the DEGs were annotated to the GO and KEGG databases. The gene temporal expression analysis during pecan kernel development was performed by STEM tool.[5] The DEGs associated with "flavonoid biosynthesis" and "phenylalanine metabolism" pathways were selected and mainly analyzed.

TABLE 1 | Summary of the RNA-Seq results.

Sample	Raw data (Gb)	Clean data (Gb)	Q30	Unique mapped	Total mapped
L100_1	6.84	6.78	94.48%	91.79%	94.69%
L100_2	6.43	6.39	94.10%	92.15%	94.89%
L100_3	7.90	7.83	94.45%	92.14%	94.93%
L114_1	8.79	8.73	93.72%	90.91%	94.33%
L114_2	7.48	7.42	93.64%	91.67%	94.69%
L114_3	10.34	10.26	94.59%	88.52%	94.88%
L121_1	7.82	7.79	94.48%	91.84%	94.42%
L121_2	6.93	6.89	95.21%	92.02%	94.53%
L121_3	8.02	7.98	95.37%	92.22%	94.78%
L135_1	7.41	7.35	94.17%	91.53%	94.67%
L135_2	8.87	8.81	94.13%	90.87%	94.77%
L135_3	7.55	7.49	94.24%	91.12%	94.53%
L150_1	8.12	8.06	94.11%	89.50%	94.28%
L150_2	6.33	6.28	94.35%	90.19%	93.91%
L150_3	6.64	6.59	93.88%	90.08%	94.34%
H100_1	5.65	5.61	93.98%	92.00%	94.78%
H100_2	4.86	4.82	94.52%	92.69%	95.27%
H100_3	5.53	5.49	93.94%	91.96%	94.75%
H114_1	9.81	9.75	95.88%	92.64%	95.38%
H114_2	8.52	8.47	95.23%	92.05%	94.71%
H114_3	6.44	6.40	95.28%	92.28%	94.86%
H121_1	8.07	8.02	95.06%	92.16%	95.37%
H121_2	9.05	9.00	94.83%	92.06%	95.31%
H121_3	7.52	7.47	94.11%	91.69%	95.10%
H135_1	8.15	8.09	94.11%	91.80%	95.07%
H135_2	5.43	5.39	94.08%	91.24%	94.71%
H135_3	7.95	7.90	94.25%	91.00%	94.51%
H150_1	8.59	8.53	94.57%	91.94%	95.44%
H150_2	7.08	7.02	94.22%	91.39%	95.12%
H150_3	6.75	6.70	94.41%	91.74%	95.43%

Correlation Analysis

For the correlation between the flavonoids and the gene expression levels, a Pearson correlation analysis was performed using the corrplot package in R.[6] A co-expression analysis was displayed by Cytoscape 3.9.0 (Shannon et al., 2003).

Transcription Factors Analysis

The transcripts annotated as R2R3-MYBs were selected. Then, the amino acid sequences of *Arabidopsis thaliana* R2R3-MYBs were downloaded from the TAIR database.[7] The amino acid

[1]http://www.geneontology.org/

[2]https://www.blast2go.com

[3]http://www.genome.jp/kegg

[4]http://planttfdb.gao-lab.org/blast.php

[5]http://www.sb.cs.cmu.edu/stem/

[6]https://github.com/taiyun/corrplot

[7]https://www.arabidopsis.org/

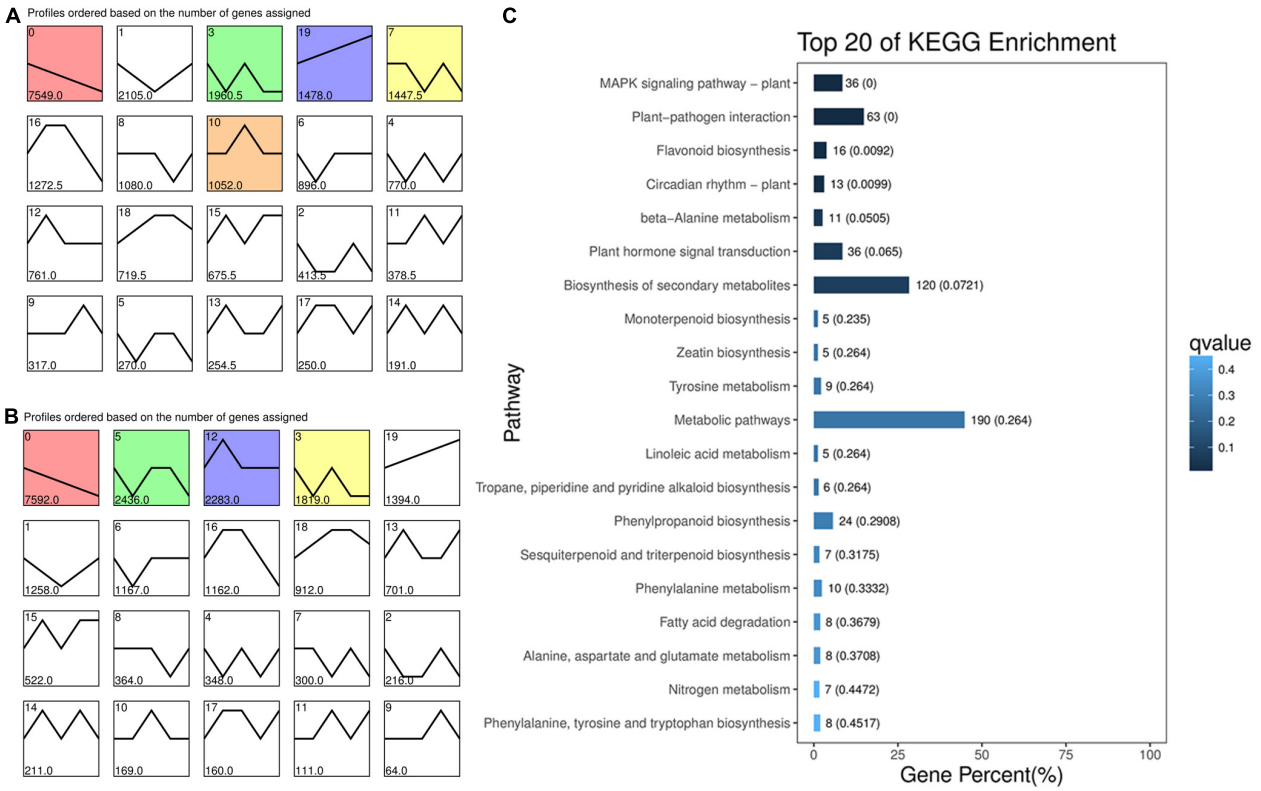


FIGURE 2 | Expression profiles of DEGs during pecan kernel development. **(A)** "Oconee"; **(B)** "YLC28"; **(C)** KEGG enrichment of the DEGs in profile 1 of "Oconee." Each square represents a gene expression profile. The number in the upper left corner indicates the profile ID number. The colored squares indicate that the profiles were significant. The horizontal axis represents five developmental stages of pecan kernels.

sequences of a pecan MYB and 120 AtMYBs were aligned using ClustalX2 software (Larkin et al., 2007). A phylogenetic tree of these sequences was constructed by MEGA 5.0[8] using a neighbor-joining method with 1000 bootstraps. Another phylogenetic tree of one pecan MYB and 11 MYBs from other plant species was generated using ClustalX2 and MEGA5.0 by the method mentioned above. A motif logo was constructed by using the online software MEME[9] and the ggseqlogo package in R (Wagih, 2017).

Quantitative Real-Time PCR Validation

To validate the accuracy of the RNA-Seq data, 14 differentially expressed genes were randomly selected for qRT-PCR (**Supplementary Table 1**). Primers were designed using Primer-Blast[10] and synthesized by TSINGKE Biotech Co., Ltd. (Hangzhou, China). Here, *18s RNA* served as the reference gene (Mattison et al., 2017). cDNA synthesis, qRT-PCR and data analysis were performed as described previously (Zhang et al., 2019a). The coefficient analysis between qRT-PCR data and RNA-Seq data was performed by using SPSS 16.0 (SPSS Inc., Chicago, IL, United States).

RESULTS AND DISCUSSION

Flavonoid Content Variation During Pecan Kernel Development

The dynamics of flavonoid along with kernel development in "YLC28" and "Oconee" were investigated (**Figure 1**). In "Oconee," the contents of TF showed an uptrend during early developmental stages and obtained the highest content of 45.56 mg/g at 128 DAFB and then decreased to 26.85 mg/g at 150 DAFB. The total flavonoid content was maintained at low levels from 100 DAFB (31.75 mg/g) to 114 DAFB (34.01 mg/g) in "YLC28" and then increased sharply from 114 DAFB to 121 DAFB (48.76 mg/g), finally reaching 1.76-fold higher levels than that of "Oconee" at 150 DAFB. In "Oconee," the CT concentration fluctuated within a narrow range from the first to the fifth stages and then decreased and reached the lowest value of 19.54 mg/g at 150 DAFB. In "YLC28," the contents of CT reached a peak on 114 DAFB (63.91 mg/g), then gradually declined and finally attained a value of 52.09 mg/g, which was 2.67-fold higher than that in "Oconee" at the same developmental stage. The contents of quercetin first increased and then decreased,

[8]www.megasoftware.net

[9]https://meme-suite.org/meme/tools/meme

[10]https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome

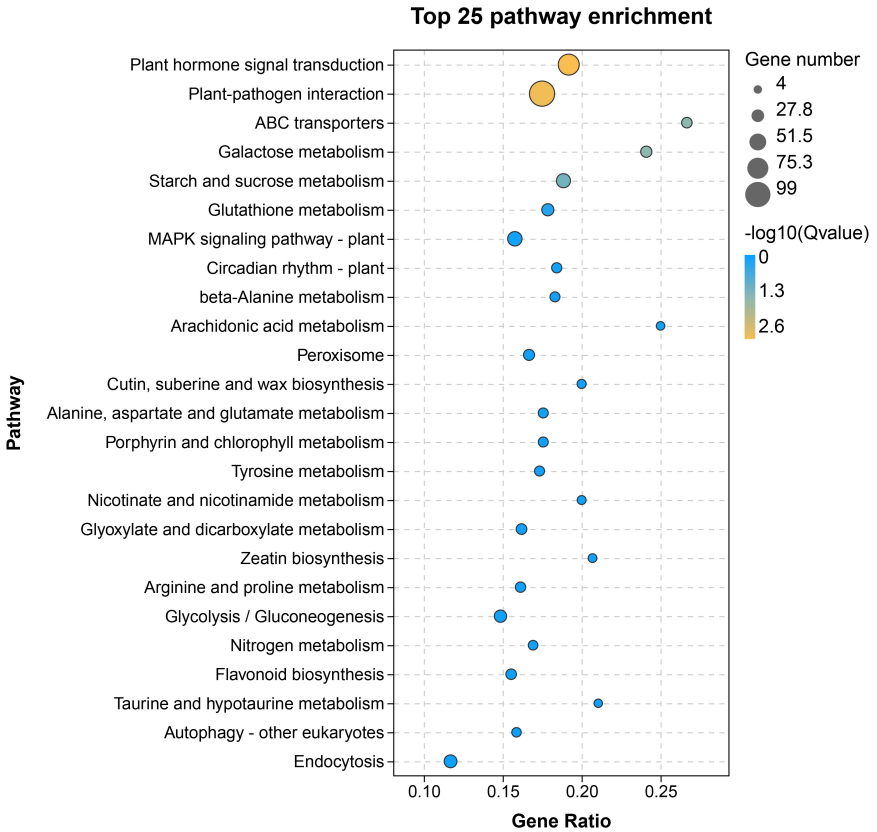


FIGURE 3 | Kyoto Encyclopedia of Genes and Genomes enrichment of upregulated genes in H121 vs. L121.

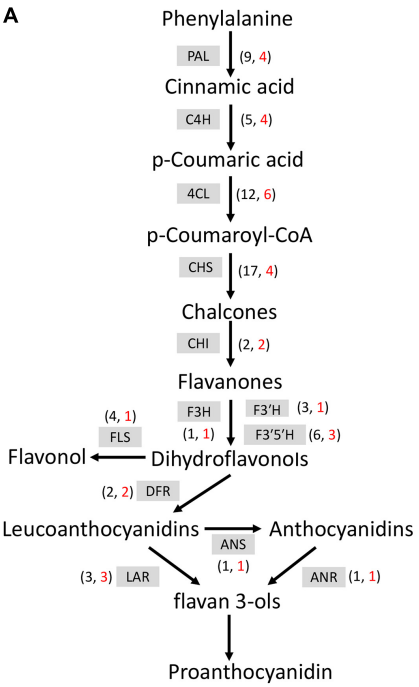
finally showing no significant difference between "YLC28" and "Oconee." The highest contents of apigenin appeared at 135 DAFB, then decreased and showed no significant difference between the two cultivars. These findings were consistent with past investigations (Jia et al., 2018; Zhang et al., 2019b), that the phenolic accumulation patterns of kernels in different genotypes were different. Here, the TF and CT contents in mature kernels of "YLC28" were 1.76- and 2.67-fold higher than that in "Oconee." Therefore, pecan kernels from the two cultivars might show different antioxidant activities and thus play different beneficial effects on human health (Venkatachalam and Sathe, 2006; Villarreal-Lozoya et al., 2007; Zhang et al., 2019b; Bouali et al., 2020). On the other hand, condensed tannins influence the seed coat color, flavor stability, and astringency in pecan nut (Polles et al., 1981). The different CT contents in nut meats might lead to different flavors and palatability between the two genotypes.

Summary of the RNA-Seq Results

A total of 224.86 Gb of raw data was generated from thirty transcriptomes (five stages with three replicates of two genotypes), and then 223.34 Gb clean data were retained after removing low-quality sequences and adaptors (**Table 1**). The average number of clean reads per sample was 49.89 million. The Q30 values of each sample were 93.64 ∼ 95.88% with an average of 94.45%. Then, all the clean reads of each sample were mapped onto the reference genome (Huang Y. et al., 2019). Approximately 93.91 ∼ 95.44% of the clean reads of each sample were successfully mapped, and 88.52 ∼ 92.69% of the clean reads were uniquely mapped. Assembled transcriptomes were annotated using 30125 reference genes (Huang Y. et al., 2019). A total of 32413 genes were obtained, including 28052 known genes and 4361 novel genes. The expression levels of each gene among 30 samples are represented by the RPKM value. A PCA analysis of all samples was performed. The results showed that the biological replicates were grouped together (**Supplementary Figure 1**), which suggested a good correlation between replicates.

Differential Gene Expression Analysis

Differential gene expression analysis was performed between the two cultivars, and along with the kernel development of each genotype. A gene temporal expression pattern analysis was carried out using a STEM software. All the DEGs in ten comparison groups, including L114 vs. L100, L121 vs. L100, L135 vs. L100, L150 vs. L100, L121 vs. L114, L135 vs. L114, L150 vs. L114, L135 vs. L121, L150 vs. L121, and L150 vs. L135, were used to generate temporal expression profiles in "Oconee" (**Figure 2A**). Similar expression profiles were also generated for "YLC28" (**Figure 2B**). All the DEGs were clustered into 20 profiles in each genotype. Most DEGs were clustered into profile



B

ID	Symbol	H100 vs L100		H114 vs L114		H121 vs L121		H135 vs L135		H150 vs L150	
		log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR
CIL0945S0098	PAL	-8.76	8E-09	8.82	6E-105	-2.78	7E-21			-6.39	5E-20
CIL0945S0099	PAL	-8.99	4E-06	8.91	1E-136	-2.98	3E-26			-5.30	2E-12
CIL1348S0047	PAL									-4.04	7E-89
CIL1196S0046	PAL			3.05	2E-239	-1.52	3E-129			-5.32	9E-77
CIL1071S0038	C4H			2.37	4E-40					-4.85	3E-09
CIL1151S0022	C4H			1.15	2E-14					-1.71	4E-09
CIL1151S0021	C4H			2.06	4E-66					-1.14	7E-08
CIL0320S0016	C4H			3.14	1E-148					-4.81	2E-08
CIL0992S0076	4CL1	-1.94	0.0422	4.59	6E-292	-1.57	4E-11			-6.62	4E-45
MSTRG.31342	4CL2	-2.88	0.0138	3.64	1E-155	-2.50	1E-89			-5.53	6E-48
CIL1192S0044	4CL3			-1.26	0.0017	1.49	8E-49			-1.38	4E-13
CIL0492S0001	4CL7	-1.12	3E-19	-1.05	5E-07						
CIL0937S0156	4CL7									-1.04	3E-08
CIL1051S0057	4CL7	-2.47	4E-07	5.13	4E-160	-2.37	3E-67	1.40	0.0039		
CIL0006S0011	CHS	-7.74	7E-06	8.43	0	-4.39	3E-241	-1.82	0.0795	-6.87	1E-18
CIL1380S0021	CHS2			-1.76	0.0004					-3.03	4E-18
CIL1399S0001	CHS2					1.61	1E-177			-2.26	1E-27
CIL1549S0032	CHS2					1.56	1E-199			-2.29	4E-24
CIL1546S0028	CHI	1.09	5E-10			2.53	5E-239			-1.68	8E-18
CIL0301S0006	CHI3					1.37	3E-122			-3.44	1E-83
CIL1020S0081	F3H									-3.14	4E-71
CIL0909S0093	F3'H			-1.32	4E-35					-3.20	3E-57
CIL0213S0013	F3'5'H					-1.30	7E-20	-1.65	0.0278	-7.44	2E-09
MSTRG.25379	F3'5'H			12.89	7E-23	-7.16	3E-07				
MSTRG.27951	F3'5'H	-9.43	0.0001	9.22	0	-3.44	5E-56			-6.67	2E-13
CIL1060S0040	FLS	-3.33	9E-16	2.52	3E-45	-3.94	2E-67				
CIL0203S0006	DFR			-2.16	0.003	1.39	8E-06				
CIL0203S0007	DFR			-2.05	8E-50	1.68	3E-263			-2.22	1E-29
CIL0526S0001	ANS			-1.11	4E-11			1.03	0.033	-2.70	5E-12
CIL1099S0044	ANR					1.35	4E-142			-2.47	7E-36
CIL0899S0128	LAR			1.42	0.0003	2.02	2E-62	2.01	7E-08	-1.18	9E-10
CIL0908S0132	LAR			-2.80	3E-17			1.93	0.063	-1.16	0.0967
CIL1032S0007	LAR	1.30	1E-09			3.28	0				

FIGURE 4 | The expression of flavonoid-related structural genes. **(A)** The flavonoid biosynthetic pathway in pecan. The numbers beside different enzymes indicate the total number (in black font) or DEG number (in red font); **(B)** The differentially expressed flavonoid-related genes during kernel development between the two genotypes.

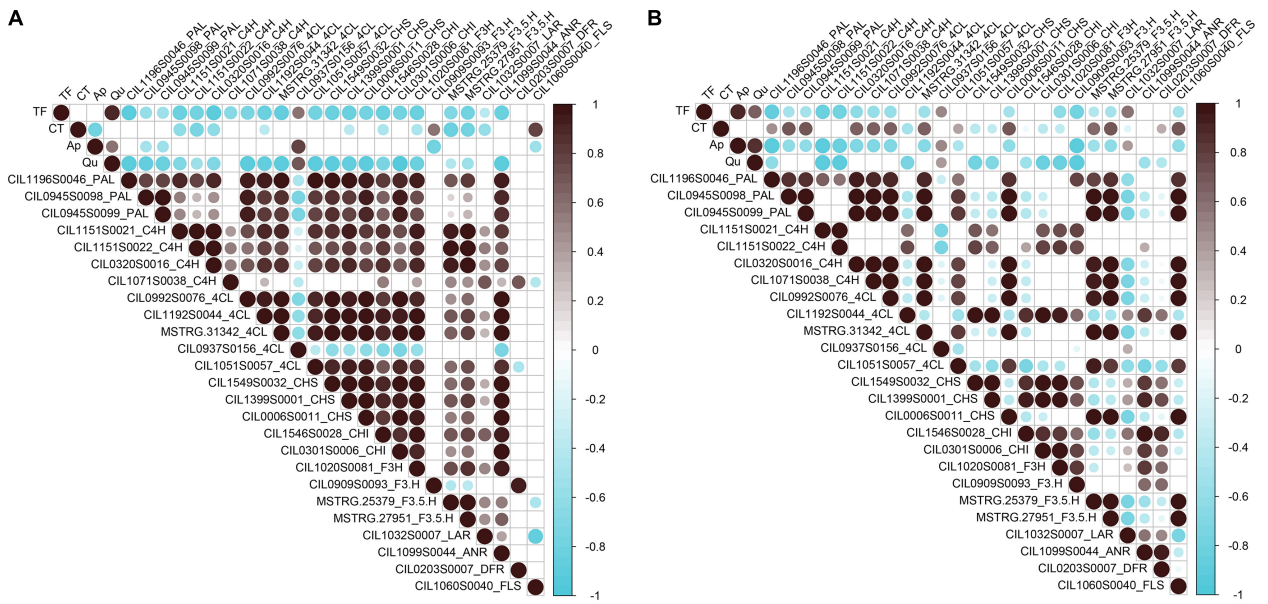


FIGURE 5 | Correlation analysis between phenolic content and structural genes. **(A)** The correlation analysis results in "Oconee"; **(B)** The correlation analysis results in "YLC28." Non-significant correlation coefficients ($p > 0.01$) are not shown.

0 in both two genotypes. The profile 0 was followed by profile 1 and profile 3 in "Oconee." However, in "YLC28," the profile 0 was followed by profile 5 and profile 12. To identify the gene functions in each profile, a KEGG pathway analysis was performed. In "Oconee," the "flavonoid biosynthesis" and "phenylalanine metabolism" pathways were significantly enriched in profile 1 (**Figure 2C**). In which, the gene expression pattern exhibited a trend of "first decreased and then rose". In total, sixteen

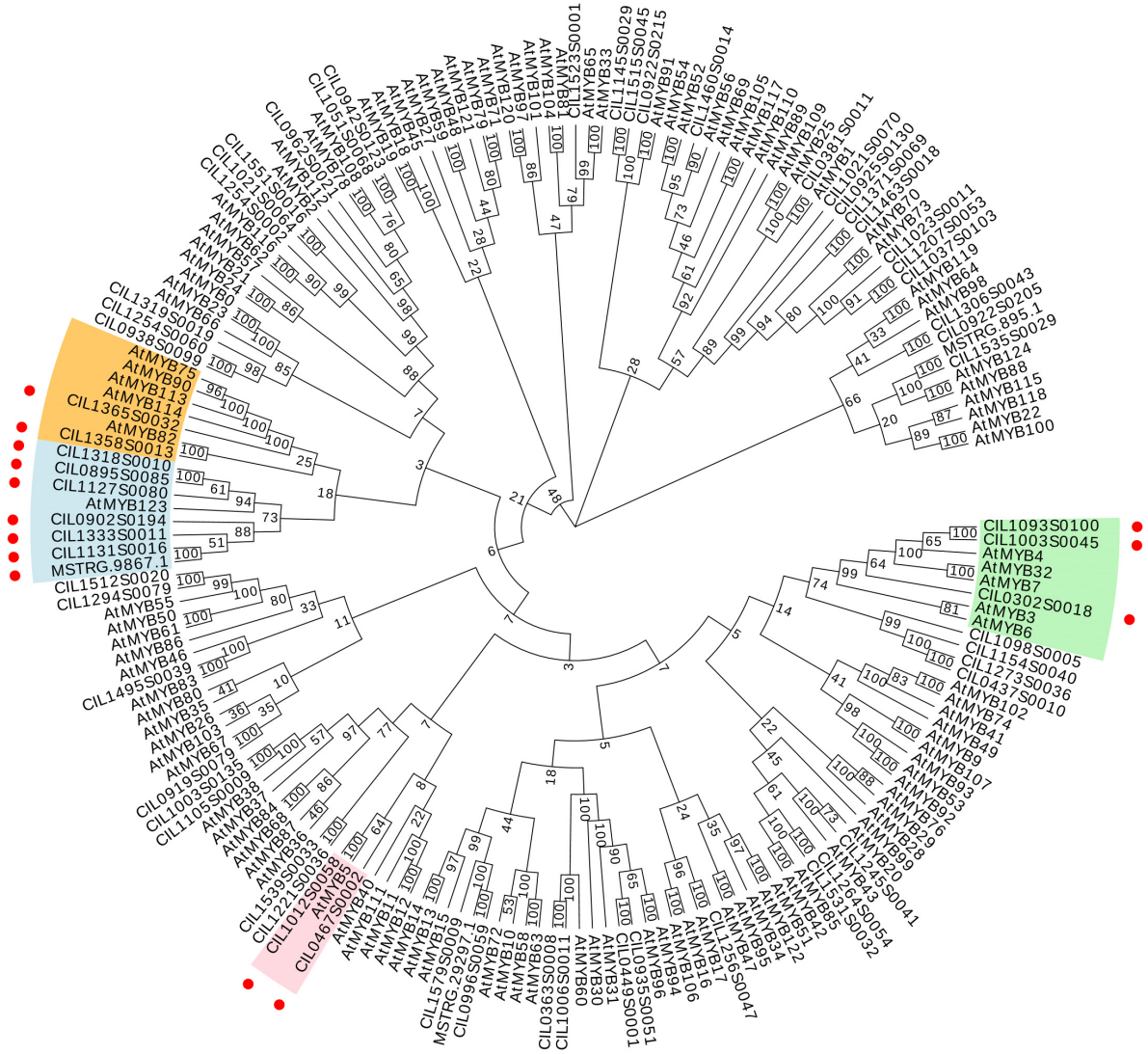


FIGURE 6 | Phylogenetic tree of 63 R2R3-MYBs from *C. illinoinensis* and 120 R2R3-MYBs from *A. thaliana*. Red circles indicate the positions of putative flavonoid-related pecan R2R3-MYBs.

and ten DEGs were enriched in "flavonoid biosynthesis" and "phenylalanine metabolism" pathways, respectively. In contrast, the expression patterns of these DEGs in "YLC28" were separately clustered in profile 0, profile 1, profile 3, profile 5, profile 6, profile 12, and profile 16 (**Supplementary Table 2**). The numbers were three, one, seven, two, one, six, and two, respectively. These results further suggested that the different expression patterns of flavonoid biosynthesis-related genes might account for the different TF and CT contents between the two genotypes.

To identify the differentially expressed flavonoid-related genes, the comparison groups of H100 vs. L100, H114 vs. L114, H121 vs. L121, H135 vs. L135, and H150 vs. L150 were analyzed. In total, 4587, 11144, 11379, 1947, and 7209 differentially expressed genes were obtained, respectively. Because the earliest significant difference of TF content between two genotypes was noticed at 121 DAFB. So, the 121 DAFB might important for the differential accumulation of flavonoids between the two cultivars. Then, the upregulated genes in "YLC28" compared to "Oconee" at 121 DAFB were selected and mapped onto KEGG pathways by the OmicShare[11] (**Figure 3**). As shown in **Figure 3**, "plant hormone signal transduction" and "plant-pathogen interaction" were the main clusters. Meanwhile, the "flavonoid biosynthesis" pathway was also significantly enriched (**Figure 3**). This result was consistent with a study in peanut that found that the "flavonoid biosynthesis" and "phenylalanine metabolism" pathways were the main categories of upregulated genes in black skin (anthocyanin-rich) compared with white skin (Huang J. et al., 2019). Similar results were also reported in walnut, in which a set of anthocyanin-related structural genes were upregulated in the leaves from a red genotype compared

[11]https://www.omicshare.com/tools/

A

ID	Symbol	H100 vs L100		H114 vs L114		H121 vs L121		H135 vs L135		H150 vs L150	
		log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR	log2(fc)	FDR
CIL0895S0085	MYB330	-7.02	1E-05	3.06	8E-30	-4.23	3E-41			-5.19	1.7E-14
CIL1358S0013	MYB6					-2.88	4E-61				
CIL1093S0100	MYB111	-1.46	7E-11	-1.11	0.0042	-2.17	0.0014	-2.21	8E-06	-5.62	4.4E-09
CIL0437S0010	C1			-1.80	2E-20	1.03	7E-10	-1.86	6E-45	-5.55	1.4E-40
MSTRG.9867	MYB5			-3.46	3E-32	1.50	4E-07				
CIL1273S0036	MYB308			-2.31	1E-07	1.55	7E-08			-2.09	0.00219
CIL1131S0016	MYB5			-1.69	2E-11	1.61	1E-27			-3.68	5.4E-11
CIL0467S0002	MYB308			-2.75	4E-101	2.08	2E-295			-1.42	1.3E-08
CIL1003S0045	MYB3	1.90	4E-15	-1.48	2E-07	2.09	3E-55			-2.91	1.4E-06
CIL1127S0080	C1					2.10	2E-15			-2.21	9.7E-06
CIL1318S0010	C1	-4.12	2E-09	1.45	3E-16	2.26	7E-27			-3.94	1.1E-33
CIL0302S0018	MYB6	1.72	0.0001			2.64	9E-74			-1.34	0.00038
CIL1154S0040	MYB4									-1.59	1E-12
CIL1012S0058	C1	1.18	6E-06	-1.01	3E-09					-2.31	1.8E-17
CIL1333S0011	TT2			-9.48	2E-05						
CIL1365S0032	MYB82	-2.00	2E-09	-5.21	5E-08						
CIL1098S0005	MYB90	1.54	2E-07								
CIL0902S0194	MYB6	1.20	2E-07	-2.79	3E-05						

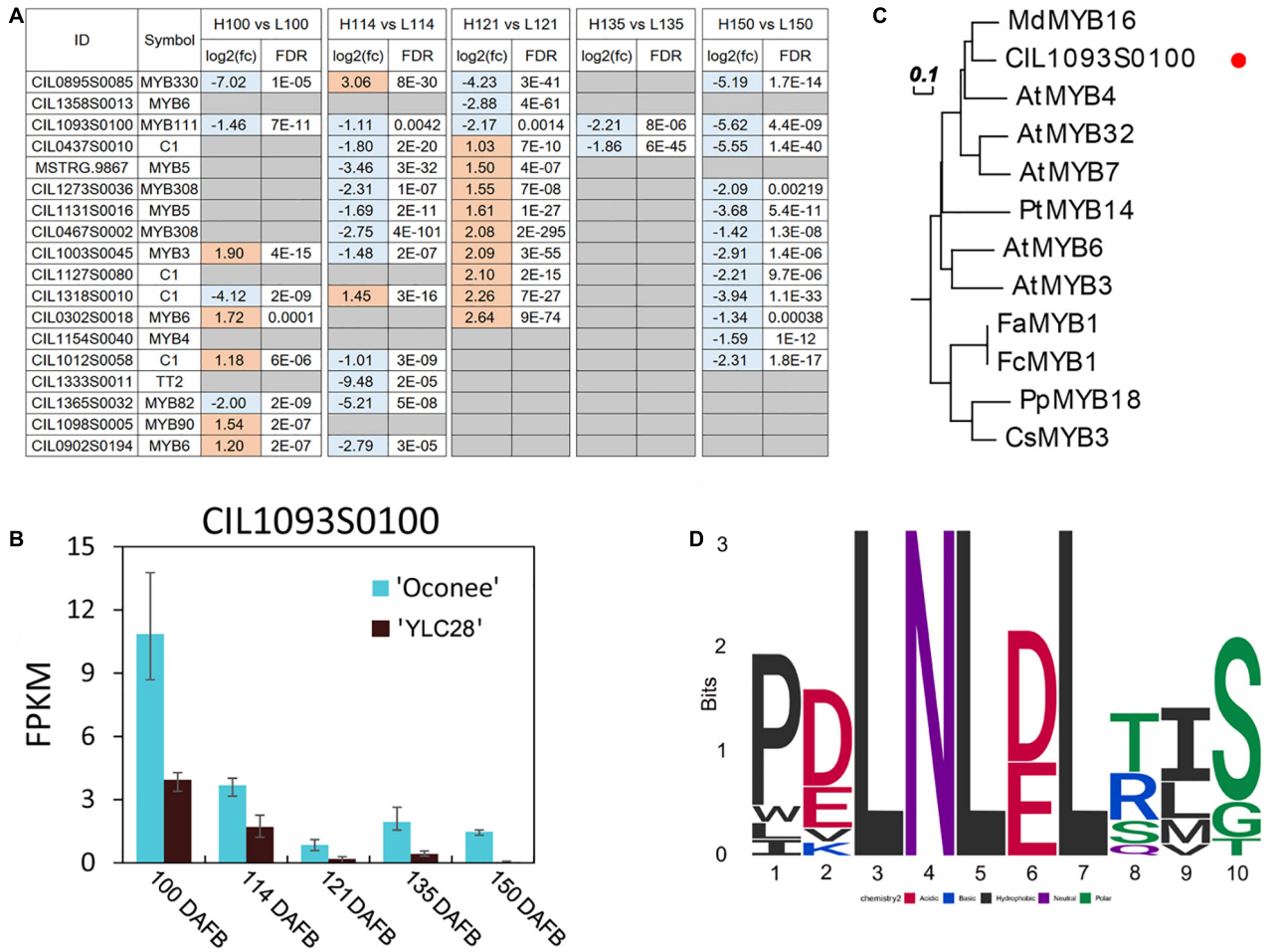


FIGURE 7 | Analysis of putative flavonoid-related pecan R2R3-MYBs. **(A)** The differentially expressed flavonoid-related *C. illinoinensis* R2R3-MYBs during kernel development between the two genotypes; **(B)** The expression of CIL1093S0100 during pecan kernel development in two cultivars. The bars indicate the maximum and minimum FPKM values; **(C)** A phylogenetic tree of subgroup 4 MYBs from pecan and other plant species; **(D)** The conserved motif of EAR (pdLNL$^{D}/_{E}$Lxi$^{G}/_{S}$) in these MYBs.

with a normal genotype (Li et al., 2018). Therefore, the high expression of flavonoid-related genes might be responsible for the higher flavonoid accumulation in "YLC28" (Li et al., 2018; Huang J. et al., 2019), and should be studied in the future.

Flavonoid Biosynthesis-Related Genes Identification

After screening the DEGs associated with flavonoid biosynthesis, 33 key flavonoid-related structure genes were obtained (**Figure 4** and **Supplementary Table 3**). These include four *PALs*, four *C4Hs*, six 4-*coumarate-CoA ligases* (*4CLs*), four *CHSs*, two *CHIs*, one *flavanone 3-hydroxylase* (*F3H*), one *flavonoid 3'-hydroxylase* (*F3'H*), three *flavonoid 3',5'-hydroxylases* (*F3'5'Hs*), two *DFRs*, three *leucoanthocyanidin reductases* (*LARs*), one *anthocyanidin synthase* (*ANS*), one *anthocyanidin reductase* (*ANR*), and one *flavonol synthase* (*FLS*). The expression trends of these genes were dramatically different between the two genotypes. In "Oconee," most of these genes exhibited relatively high expression levels at the first (100 DAFB) and fifth (150 DAFB) developmental stages. In contrast, these genes were upregulated from the first to the third stages in "YLC28." The transcription levels of three *PALs* (CIL1196S0046, CIL0945S0098, and CIL0945S0099), four *C4Hs* (CIL0320S0016, CIL1071S0038, CIL1151S0021, and CIL1151S0021), three *4CLs* (CIL1051S0057, MSTRG.31342, and CIL0992S0076), one *CHS* (CIL0006S0011), two *F3'5'Hs* (MSTRG.25379 and MSTRG.27951), and one *FLS* (CIL1060S0040) in the second stage (114 DAFB) of "YLC28" were significantly increased compared with those in "Oconee" at the same stage. One *4CL* (CIL1192S0044), two *CHSs* (CIL1399S0001 and CIL1549S0032), one *CHI* (CIL0301S0006), one *F3H* (CIL1020S0081), one *F3'H* (CIL0909S0093), and one *DFR* (CIL0203S0007) gene were upregulated at stage three (121 DAFB) in "YLC28" compared with "Oconee." The contents of CT and TF in "YLC28" were significantly increased compared with those in "Oconee" at 114 and 121 DAFB, respectively. The high expression of these structural genes might be important for abundant flavonoid accumulation in "YLC28" kernels. In

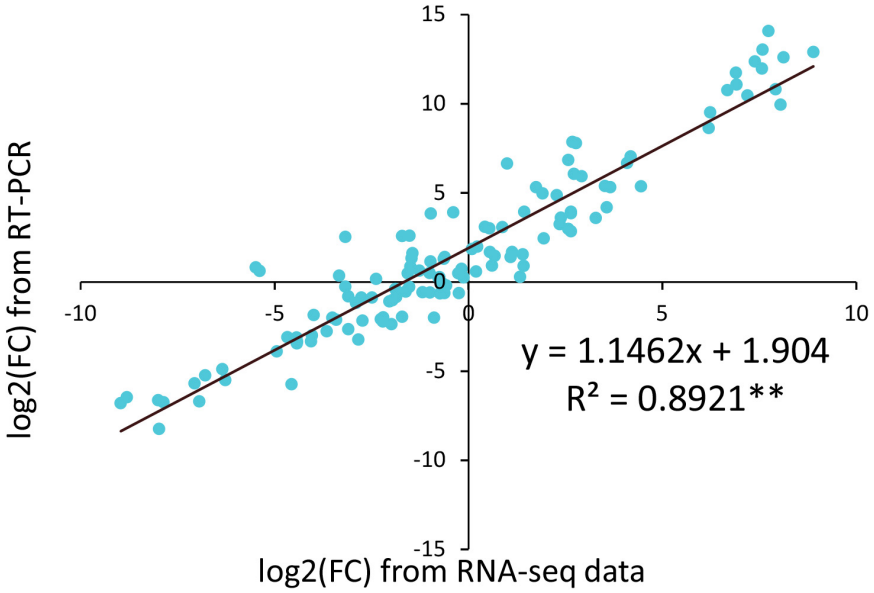


FIGURE 8 | Coefficient analysis between qRT-PCR and RNA-Seq. FC, fold change.

addition, one *4CL* (CIL0937S0156), one *CHI* (CIL1546S0028), one *LAR* (CIL1032S0007) and one *ANR* (CIL1099S0044) gene exhibited increased transcription levels in the third and fourth (135 DAFB) stages in "YLC28" compared with "Oconee." Similarly, the high expression of structural genes in organs rich in flavonoids has also been reported in potato (Liu et al., 2015), walnut (Li et al., 2018), *Camellia sinensis* (Wei et al., 2016), *Zanthoxylum bungeanum* (Sun et al., 2019), blueberry (Lin et al., 2018), and peanut (Huang J. et al., 2019).

Correlation Analysis Between Flavonoid Content and Genes

To identify key genes associated with the biosynthesis and accumulation of flavonoid components in pecan kernels, Pearson correlation analysis between the contents of the components and the flavonoid-related DEGs was performed. As shown in **Figure 5**, the correlation analysis results showed dramatic differences between "Oconee" and "YLC28." In "Oconee," only one *4CL* (CIL0937S0156) one *F3'H* (CIL0909S0093) and one *FLS* (CIL1060S0040) exhibited a significant positive correlation ($r > 0.5$, $p < 0.01$) with the contents of TF, CT, Ap, and Qu. In "YLC28," one, ten and one genes ($r > 0.5$, $p < 0.01$) exhibited positive correlations with the TF, CT, and Ap, respectively. Among them, two *PALs* (CIL0945S0098 and CIL0945S0099), one *CHS* (CIL0006S0011), one *F3'5'H* (MSTRG.27951), and one *FLS* (CIL1060S0040) exhibited high correlation ($r \geq 0.7$, $p < 0.01$) with the CT content in "YLC28." They might play crucial roles in the CT biosynthesis in "YLC28." The content of CT can affect the flavor and palatability of pecan nut (Polles et al., 1981). Thus, these genes will be useful for the CT trait improvement in pecan. In addition, co-expression between structural genes was also noticed (**Figure 5** and **Supplementary Figure 2**). Twenty-four genes, including three *PALs*, four *C4Hs*, four *4CLs*, three *CHSs*, two *CHIs*, two *F3'5'Hs*, one *F3H*, one *F3'H*, one *FLS*, one *ANR*, one *DFR*, and one *LAR*, shared strong positive correlations (Pearson's $r \geq 0.65$, $p < 0.05$). Similar results were reported in strawberry (Pillet et al., 2015) and wintersweet flower (Yang et al., 2018). In strawberry, the expression of thirteen flavonoid-related genes showed positive correlations ($r > 0.65$, $P < 0.05$) with each other (Pillet et al., 2015). Yang et al. (2018) reported that the transcription of *F3H1*, *F3'H1*, *ANS1*, *UFGT1*, and *CHSs* shared strong correlations in wintersweet flower. These results indicated that the remarkably different accumulation profiles of flavonoids between "YLC28" and "Oconee" resulted from the distinct gene expression patterns. Furthermore, flavonoid-related genes are closely connected and may be controlled by the same transcription factors (Yang et al., 2018).

Flavonoid Related Transcription Factors Identification

To predict flavonoid-related transcription factors in pecan kernels, all the genes were annotated according to the PlantTFdb. In total, 926 genes were predicted as transcription factors. Among the different types of transcription factors, bHLH (83), ERF (74), MYB (70), NAC (54), and MYB_related (45) were the main categories (**Supplementary Figure 3**). R2R3-MYB transcription factors regulate the transcription of flavonoid biosynthetic genes in various plants (Xu et al., 2017). All the potential pecan MYBs were analyzed, and a total of 63 R2R3-MYBs were obtained. Then, a phylogenetic tree of pecan MYBs with AtMYBs was constructed (**Figure 6**). Previous studies reported that AtMYB75, AtMYB90, AtMYB113, and AtMYB114 are involved in the regulation of anthocyanin biosynthesis (Dubos et al., 2010). AtMYB123 and AtMYB5 regulate the accumulation of proanthocyanidin

in *Arabidopsis* (Nesi et al., 2001; Schaart et al., 2013). AtMYB3, AtMYB4, AtMYB6, AtMYB7, and AtMYB32 act as repressors that control phenylpropanoid and flavonoid biosynthesis (Ma and Constabel, 2019). As shown in **Figure 6**, fourteen pecan MYBs clustered together with the AtMYBs mentioned above. Therefore, these MYB sequences might be associated with flavonoid accumulation during pecan kernel maturation.

Subsequently, the gene differential expression analysis result of 14 *MYBs* was illustrated in **Figure 7A**. They exhibited distinct expression patterns between the two cultivars, which suggested that the MYBs might exhibit subtle control of structural genes. Among them, the expression levels of CIL1093S0100 were significantly upregulated in "Oconee" at all developmental stages than those in "YLC28" (**Figures 7A,B**). The predicted amino acid of CIL1093S0100 shared high identity with the subgroup 4 MYB member AtMYB4 (identity = 53.37%). AtMYB4 functions as a repressor in phenylpropanoid biosynthesis by negatively regulating the synthesis of sinapate esters by repressing the transcription of C4H (Jin et al., 2000). Subgroup 4 MYBs contain a conserved ethylene-responsive element-binding factor-associated amphiphilic repression (EAR) motif, which directly binds to the promoters of target genes and acts as repressors of the flavonoid or lignin pathway (Wan et al., 2017). A phylogenetic tree was constructed by CIL1093S0100 and 11 members of the subgroup 4 MYBs from other species (**Figure 7C**). CIL1093S0100 clustered together with AtMYB4 and shared high similarity with MdMYB16 (identity = 64.43%). A multiple sequence alignment of these sequences was generated (**Supplementary Figure 4**). The typical R2 and R3 domains of the R2R3-MYB transcription factors and a conserved motif EAR (pdLNLD/ELxiG/S) were found in these MYBs (**Supplementary Figure 4** and **Figure 7D**). In poplar, overexpression or knockout of the AtMYB4-like gene *PtrMYB57* led to reduced or increased anthocyanin and CT contents, respectively (Wan et al., 2017). Similarly, CsMYB3 and PpMYB18 act as negative regulators of anthocyanin and CT accumulation in citrus and peach, respectively (Huang D. et al., 2019; Zhou et al., 2019). In apple, MdMYB16 regulates the expression of MdANS and MdUFGT, and overexpression of *MdMYB16* in the red-fleshed callus resulted in a decrease in anthocyanin content (Dubos et al., 2010). Because CIL1093S0100 shares a high identity with MdMYB16, the two transcription factors might perform similar functions. Recently, the expression profiles of *R2R3-MYBs* have been detected during graft union formation in pecan. The CIL1093S0100 showed an upregulated trend along with the graft union formation (Mo et al., 2018). On the contrary, the transcription levels of four *CiPALs* were downregulated during this physiological process (Zhang et al., 2022). Previous study reported that the increase of phenolic and flavonoid content was associated with the graft-incompatibility between rootstock and scion (Zhu et al., 2020). Therefore, the higher expression levels of CIL1093S0100 might be related to lower contents of flavonoids and condensed tannins in pecan. The identification of target genes and the regulatory mechanism of CIL1093S0100 should be performed in the near future.

RNA-Seq Validation by Quantitative Real-Time PCR

To verify the RNA-Seq data, the expression of fourteen randomly selected genes was analyzed using qRT-PCR. The qRT-PCR results of most genes were consistent with the RNA-Seq data (**Supplementary Figure 5**). Then, a linear regression analysis was performed using RNA-Seq data and qRT-PCR results (**Figure 8**). A high correlation coefficient of 0.8921 ($p < 0.01$) was observed, indicating that the transcriptome data were reliable.

Flavonoids are key nutrition components for pecan nuts. However, the molecular mechanisms of flavonoid biosynthesis in pecan kernels have not been fully elucidated. Here, the flavonoid content changes and gene expression patterns during embryo development in "YLC28" and "Oconee" were detected. The contents changing patterns of TF and CT were significantly different between the two cultivars during kernel development. RNA-Seq results indicated that the upregulated genes in "YLC28" were significantly enriched in flavonoid biosynthesis pathways. Thirty-three differentially expressed flavonoid-related structural genes were obtained. The expression profiles of two *PALs*, one *CHS*, one *F3'5'H*, and one *FLS* were significantly correlated with the CT content changes in "YLC28." A putative subgroup 4 MYB transcription factor (CIL1093S0100) was identified, which might act as a flavonoid biosynthesis repressor during kernel development. The present study will be useful for accelerating the study of the molecular basis of flavonoid biosynthesis in pecan kernels.

DATA AVAILABILITY STATEMENT

The raw data have been deposited in the NCBI database with the bioproject accession number PRJNA792564.

AUTHOR CONTRIBUTIONS

CZ, XY, and HR conceptualized the project. CZ performed the experiments, did the data analysis, and wrote the manuscript. KW and JC helped CZ to did the data analysis. All authors reviewed and approved the manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China No. 31800575, and the Fundamental Research Funds of CAF Nos. CAFYBB2018SY013 and CAFYBB2017ZA004-8.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.804968/full#supplementary-material

Supplementary Figure 1 | The PCA analysis of all samples.

Supplementary Figure 2 | The co-expression analysis of flavonoid-related DEGs. The size of the node indicates the connectivity degree. The width of edge denotes the strength of correlation.

Supplementary Figure 3 | The number of different types of transcription factors.

Supplementary Figure 4 | Multiple alignment of amino acid sequences of CIL1093S0100 with 11 subgroup 4 MYBs from other species.

Supplementary Figure 5 | Expression patterns of 14 randomly selected genes. The column diagrams and line charts represent the RT-PCR results and the RPKM values of genes. The abscissa axis indicates five different stages of kernels. The error bars indicate standard deviation of three replicates.

Supplementary Table 1 | Primers used in this study.

Supplementary Table 2 | Genes related to phenylalanine metabolism and flavonoid biosynthesis pathways in profile 1 of "Oconee."

Supplementary Table 3 | The information of flavonoid-related genes.

REFERENCES

Akagi, T., Ikegami, A., Tsujimoto, T., Kobayashi, S., Sato, A., Kono, A., et al. (2009). DkMyb4 is a myb transcription factor involved in proanthocyanidin biosynthesis in persimmon fruit. *Plant Physiol.* 151, 2028–2045. doi: 10.1104/pp.109.146985

Bouali, I., Tsafouros, A., Ntanos, E., Albouchi, A., Boukhchina, S., and Roussos, P. A. (2020). Inter-cultivar and temporal variation of phenolic compounds, antioxidant activity and carbohydrate composition of pecan (*Carya illlinoinensis*) kernels grown in Tunisia. *Horticul. Environ. Biotechnol.* 61, 183–196.

Caltagirone, S., Rossi, C., Poggi, A., Ranelletti, F. O., Natali, P. G., Brunetti, M., et al. (2000). Flavonoids apigenin and quercetin inhibit melanoma growth and metastatic potential. *Int. J. Cancer* 87, 595–600. doi: 10.1002/1097-0215(20000815)87:4<595::aid-ijc21>3.0.co;2-5

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

de la Rosa, L. A., Alvarez-Parrilla, E., and Shahidi, F. (2011). Phenolic compounds and antioxidant activity of kernels and shells of mexican pecan (*Carya illinoinensis*). *J. Agr. Food Chem.* 59, 152–162. doi: 10.1021/jf1034306

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in arabidopsis. *Trends Plant Sci.* 15, 573–581.

Gong, Y., and Pegg, R. B. (2017). Separation of ellagitannin-rich phenolics from U.S. pecans and chinese hickory nuts using fused-core HPLC columns and their characterization. *J. Agr. Food Chem.* 65, 5810–5820. doi: 10.1021/acs.jafc.7b01597

Grauke, L. J., Wood, B. W., and Harris, M. K. (2016). Crop vulnerability: carya. *J. Am. Soc. Hortic. Sci.* 51:653. doi: 10.21273/hortsci.51.6.653

Harnly, J. M., Doherty, R. F., Beecher, G. R., Holden, J. M., Haytowitz, D. B., Bhagwat, S., et al. (2006). Flavonoid content of U.S. fruits, vegetables, and nuts. *J. Agr. Food Chem.* 54, 9966–9977. doi: 10.1021/jf061478a

Huang, D., Tang, Z., Fu, J., Yuan, Y., Deng, X., and Xu, Q. (2019). CsMYB3 and CsRuby1 form an 'activator-and-repressor' loop for the regulation of anthocyanin biosynthesis in citrus. *Plant Cell Physiol.* 61, 318–330. doi: 10.1093/pcp/pcz198

Huang, J., Xing, M., Li, Y., Cheng, F., Gu, H., Yue, C., et al. (2019). Comparative transcriptome analysis of the skin-specific accumulation of anthocyanins in black peanut (*Arachis hypogaea* L.). *J. Agr. Food Chem.* 67, 1312–1324. doi: 10.1021/acs.jafc.8b05915

Huang, Y., Xiao, L., Zhang, Z., Zhang, R., Wang, Z., Huang, C., et al. (2019). The genomes of pecan and Chinese hickory provide insights into carya evolution and nut nutrition. *GigaScience* 8:giz036.

Jia, X., Luo, H., Xu, M., Zhai, M., Guo, Z., Qiao, Y., et al. (2018). Dynamic changes in phenolics and antioxidant capacity during pecan (*Carya illinoinensis*) kernel ripening and its phenolics profiles. *Molecules* 23:435. doi: 10.3390/molecules23020435

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

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kornsteiner, M., Wagner, K.-H., and Elmadfa, I. (2006). Tocopherols and total phenolics in 10 different nut types. *Food Chem.* 98, 381–387. doi: 10.1016/j.foodchem.2005.07.033

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Li, Y., Luo, X., Wu, C., Cao, S., Zhou, Y., Jie, B., et al. (2017). Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in red and green walnut *(Juglans regia L.)*. *Molecules* 23:25. doi: 10.3390/molecules23010025

Li, Y., Luo, X., Wu, C., Cao, S., Zhou, Y., Jie, B., et al. (2018). Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in red and green walnut (*Juglans regia* L.). *Molecules* 23:25. doi: 10.3390/molecules23010025

Lin, Y., Wang, Y., Li, B., Tan, H., Li, D., Li, L., et al. (2018). Comparative transcriptome analysis of genes involved in anthocyanin synthesis in blueberry. *Plant Physiol. Bioch.* 127, 561–572. doi: 10.1016/j.plaphy.2018.04.034

Liu, Y., Lin-Wang, K., Deng, C., Warran, B., Wang, L., Yu, B., et al. (2015). Comparative transcriptome analysis of white and purple potato to identify genes involved in anthocyanin biosynthesis. *PLos One* 10:e0129148. doi: 10.1371/journal.pone.0129148

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNAseq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Ma, D., and Constabel, C. P. (2019). MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends Plant Sci.* 24, 275–289. doi: 10.1016/j.tplants.2018.12.003

Mattison, C. P., Rai, R., Settlage, R. E., Hinchliffe, D. J., Madison, C., Bland, J. M., et al. (2017). RNA-seq analysis of developing pecan (*Carya illinoinensis*) embryos reveals parallel expression patterns among allergen and lipid metabolism genes. *J. Agr. Food Chem.* 65, 1443–1455. doi: 10.1021/acs.jafc.6b04199

Mo, Z., Feng, G., Su, W., Liu, Z., and Peng, F. (2018). Transcriptomic analysis provides insights into grafting union development in pecan (*Carya illinoinensis*). *Genes* 9:71. doi: 10.3390/genes9020071

Nesi, N., Jond, C., Debeaujon, I., Caboche, M., and Lepiniec, L. (2001). The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* 13, 2099–2114. doi: 10.1105/tpc.010098

Pandey, K. B., and Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxid. Med. Cell. Longev.* 2, 270–278.

Pérez-Jiménez, J., Neveu, V., Vos, F., and Scalbert, A. (2010). Identification of the 100 richest dietary sources of polyphenols: an application of the phenol-explorer database. *Eur. J. Clin.l Nutrition* 64:S112. doi: 10.1038/ejcn.2010.221

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Pillet, J., Yu, H.-W., Chambers, A. H., Whitaker, V. M., and Folta, K. M. (2015). Identification of candidate flavonoid pathway genes using transcriptome correlation network analysis in ripe strawberry (fragaria × ananassa) fruits. *J. Exp. Bot.* 66, 4455–4467. doi: 10.1093/jxb/erv205

Polles, S. G., Hanny, B. W., and Harvey, A. J. (1981). Condensed tannins in kernels of thirty-one pecan [*Carya illinoensis* (wangenh) K. koch] cultivars. *J. Agr. Food Chem.* 29, 196–197. doi: 10.1021/jf00103a052

Robbins, K. S., Ma, Y., Wells, M. L., Greenspan, P., and Pegg, R. B. (2014). Separation and characterization of phenolic compounds from U.S. pecans by

liquid chromatography–tandem mass spectrometry. *J. Agr. Food Chem.* 62, 4332–4341. doi: 10.1021/jf500909h

Schaart, J. G., Dubos, C., Romero De La Fuente, I., van Houwelingen, A. M. M. L., de Vos, R. C. H., Jonker, H. H., et al. (2013). Identification and characterization of MYB-bHLH-WD40 regulatory complexes controlling proanthocyanidin biosynthesis in strawberry (fragaria × ananassa) fruits. *New Phytol.* 197, 454–467. doi: 10.1111/nph.12017

Senter, S. D., Horvat, R. J., and Forbus, W. R. Jr. (1980). Relation between phenolic acid content and stability of pecans in accelerated storage. *J. Food Sci.* 45, 1380–1382. doi: 10.1111/j.1365-2621.1980.tb06559.x

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504.

Shao, Z., Zhang, P., Lu, C., Li, S., Chen, Z., Wang, X., et al. (2019). Transcriptome sequencing of *Saccharina japonica* sporophytes during whole developmental periods reveals regulatory networks underlying alginate and mannitol biosynthesis. *BMC Genomics* 20:975. doi: 10.1186/s12864-019-6366-x

Shay, J., Elbaz, H. A., Lee, I., Zielske, S. P., Malek, M. H., and Hüttemann, M. (2015). Molecular mechanisms and therapeutic effects of (-)-epicatechin and other polyphenols in cancer, inflammation, diabetes, and neurodegeneration. *Oxid. Med. Cell. Longev.* 2015:181260. doi: 10.1155/2015/181260

Sun, L., Yu, D., Wu, Z., Wang, C., Yu, L., Wei, A., et al. (2019). Comparative transcriptome analysis and expression of genes reveal the biosynthesis and accumulation patterns of key flavonoids in different varieties of *Zanthoxylum bungeanum* leaves. *J. Agr. Food Chem.* 67, 13258–13268. doi: 10.1021/acs.jafc.9b05732

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

Venkatachalam, M., and Sathe, S. K. (2006). Chemical composition of selected edible nut seeds. *J. Agr. Food Chem.* 54, 4705–4714. doi: 10.1021/jf0606959

Villarreal-Lozoya, J. E., Lombardini, L., and Cisneros-Zevallos, L. (2007). Phytochemical constituents and antioxidant capacity of different pecan [*Carya illinoinensis* (wangenh.) K. koch] cultivars. *Food Chem.* 102, 1241–1249.

Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* 33, 3645–3647. doi: 10.1093/bioinformatics/btx469

Wan, S., Li, C., Ma, X., and Luo, K. (2017). PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. *Plant Cell Rep.* 36, 1263–1276. doi: 10.1007/s00299-017-2151-y

Wang, Y., Gao, Y., Ding, H., Liu, S., Han, X., Gui, J., et al. (2017). Subcritical ethanol extraction of flavonoids from Moringa oleifera leaf and evaluation of antioxidant activity. *Food Chem.* 218, 152–158. doi: 10.1016/j.foodchem.2016.09.058

Wei, K., Zhang, Y., Wu, L., Li, H., Ruan, L., Bai, P., et al. (2016). Gene expression analysis of bud and leaf color in tea. *Plant Physiol. Bio.* 107, 310–318. doi: 10.1016/j.plaphy.2016.06.022

Xu, H., Wang, N., Liu, J., Qu, C., Wang, Y., Jiang, S., et al. (2017). The molecular mechanism underlying anthocyanin metabolism in apple using the MdMYB16 and MdbHLH33 genes. *Plant Mol. Biol.* 94, 149–165. doi: 10.1007/s11103-017-0601-0

Yang, N., Zhao, K., Li, X., Zhao, R., Aslam, M. Z., Yu, L., et al. (2018). Comprehensive analysis of wintersweet flower reveals key structural genes involved in flavonoid biosynthetic pathway. *Gene* 676, 279–289. doi: 10.1016/j.gene.2018.08.050

Zhang, C., Ren, H., Yao, X., Wang, K., and Chang, J. (2021). Full-length transcriptome analysis of pecan (*Carya illinoinensis*) kernels. *G3 Genes Genom. Genet.* 11:jkab182.

Zhang, C., Yao, X., Ren, H., Chang, J., and Wang, K. (2019a). RNA-seq reveals flavonoid biosynthesis-related genes in pecan (*Carya illinoinensis*) kernels. *J. Agr. Food Chem.* 67, 148–158. doi: 10.1021/acs.jafc.8b05239

Zhang, C., Yao, X., Ren, H., Wang, K., and Chang, J. (2019b). Isolation and characterization of three chalcone synthase genes in pecan (*Carya illinoinensis*). *Biomolecules* 9:236. doi: 10.3390/biom9060236

Zhang, C., Yao, X., Ren, H., Wang, K., and Chang, J. (2022). Genome-wide identification and characterization of the phenylalanine ammonia-lyase gene family in pecan (*Carya illinoinensis*). *Sci. Hortic.* 295:110800.

Zhou, H., Lin-Wang, K., Wang, F., Espley, R. V., Ren, F., Zhao, J., et al. (2019). Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and proanthocyanidin accumulation. *New Phytol.* 221, 1919–1934. doi: 10.1111/nph.15486

Zhu, K., Fan, P., Mo, Z., Tan, P., Feng, G., Li, F., et al. (2020). Identification, expression and co-expression analysis of R2R3-MYB family genes involved in graft union formation in pecan (*Carya illinoinensis*). *Forests* 11:917.

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 March 2022
doi: 10.3389/fpls.2022.849048

Function Analysis of the ERF and DREB Subfamilies in Tomato Fruit Development and Ripening

Li Zhang[2,5], LiJing Chen[2,5], ShengQun Pang[1,4], Qun Zheng[1,4], ShaoWen Quan[1,4], YuFeng Liu[3], Tao Xu[3], YuDong Liu[1,4] and MingFang Qi[3]**

[1] College of Agriculture, Shihezi University, Shihezi, China, [2] College of Bioscience and Biotechnology, Shenyang Agricultural University, Shenyang, China, [3] College of Horticulture, Shenyang Agricultural University, Shenyang, China, [4] Key Laboratory of Special Fruits and Vegetables Cultivation Physiology and Germplasm Resources Utilization Xinjiang of Production and Construction Crops, Shihezi University, Shihezi, China, [5] Key Laboratory of Agricultural Biotechnology of Liaoning Province, Shenyang Agricultural University, Shenyang, China

OPEN ACCESS

Edited by:
Shunquan Lin, South China Agricultural University, China

Reviewed by:
Sixue Chen, University of Florida, United States
Li Zhengguo, Chongqing University, China

***Correspondence:**
YuDong Liu
lyd-forever@163.com
MingFang Qi
qimingfang@syau.edu.cn

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

Received: *05 January 2022*
Accepted: *02 February 2022*
Published: *04 March 2022*

Citation:
Zhang L, Chen L, Pang S, Zheng Q, Quan S, Liu Y, Xu T, Liu Y and Qi M (2022) Function Analysis of the ERF and DREB Subfamilies in Tomato Fruit Development and Ripening. Front. Plant Sci. 13:849048. doi: 10.3389/fpls.2022.849048

APETALA2/ethylene responsive factors (AP2/ERF) are unique regulators in the plant kingdom and are involved in the whole life activity processes such as development, ripening, and biotic and abiotic stresses. In tomato (*Solanum lycopersicum*), there are 140 AP2/ERF genes; however, their functionality remains poorly understood. In this work, the 14th and 19th amino acid differences in the AP2 domain were used to distinguish DREB and ERF subfamily members. Even when the AP2 domain of 68 ERF proteins from 20 plant species and motifs in tomato DREB and ERF proteins were compared, the binding ability of DREB and ERF proteins with DRE/CRT and/or GCC boxes remained unknown. During fruit development and ripening, the expressions of 13 DREB and 19 ERF subfamily genes showed some regular changes, and the promoters of most genes had ARF, DRE/CRT, and/or GCC boxes. This suggests that these genes directly or indirectly respond to IAA and/or ethylene (ET) signals during fruit development and ripening. Moreover, some of these may feedback regulate IAA or ET biosynthesis. In addition, 16 EAR motif-containing ERF genes in tomato were expressed in many organs and their total transcripts per million (TPM) values exceeded those of other ERF genes in most organs. To determine whether the EAR motif in EAR motif-containing ERF proteins has repression function, their EAR motifs were retained or deleted in a yeast one-hybrid (YIH) assay. The results indicate that most of EAR motif-containing ERF proteins lost repression activity after deleting the EAR motif. Moreover, some of these were expressed during ripening. Thus, these EAR motif-containing ERF proteins play vital roles in balancing the regulatory functions of other ERF proteins by completing the DRE/CRT and/or GCC box sites of target genes to ensure normal growth and development in tomato.

Keywords: AP2/ERF, tomato (*Solanum lycopersicum*), ERF, DREB, DRE/CRT, GCC box, EAR motif, yeast one-hybrid

INTRODUCTION

Plant hormones are involved in vital processes of complex signal transduction pathways and affect the expression of various genes at different time periods and in different organs, which regulate plant growth, development, and defense responses. To ensure survival and reproduction, diverse hormones, such as auxin (IAA), abscisic acid (ABA), ethylene (ET), gibberellin (GA), cytokinin (CTK), and jasmonate (JA), are synthetized and regulate different life activities in their metabolic networks (Davies, 1987; Pieterse et al., 2009). In these complex networks, transcription factors (TFs) are critical regulators that play essential roles (Yamasaki et al., 2013).

Among these TFs, AP2/ERF is widely distributed in the plant kingdom and plays important roles in regulating growth and development (Chen H. C. et al., 2021; Guo et al., 2021; Jia et al., 2021). With updates to the plant genome database, more AP2/ERF genes have been identified. Thus far, 147, 291, 170, 163, and 136 AP2/ERF genes have been found in *Arabidopsis thaliana* (*A. thaliana*) (Dietz et al., 2010), *Chinese cabbage* (Song et al., 2013), *Salvia miltiorrhiza* (Ji et al., 2016), rice (Akhter Most et al., 2011), and melon (Ma et al., 2015). Currently, RNA sequencing of many species has been conducting, laying an important foundation for the study of AP2/ERF gene families during plant growth and development. *Jatropha curcas* L. *JcERF035* was identified in the roots and leaves under Pi deficiency conditions by RNA sequencing and its overexpression affected root development in *A. thaliana* (Chen et al., 2018).

AP2/ERF family members have been divided into 5 groups according to *A. thaliana* classifications: ERF, DREB, AP2, RAVs, and soloist (Nakano et al., 2006; Dietz et al., 2010). These subfamilies exhibit different structural characteristics. Among them, differences between ERF and DREB subfamily members is that the 14th and 19th amino acids of the DREB proteins are valine (V) and glutamate (E) in the AP2 domain, but alanine (A) and aspartate (D) in ERF proteins. This difference affects the ability of proteins to interact with DRE or GCC boxes during the regulation of their downstream target genes during transcription (Sakuma et al., 2002). It also suggests that ERF and DREB subfamily members may act in different regulatory pathways. For example, *PUCHI*, an ERF subfamily gene, regulates lateral root development, floral meristem identity, and organ initiation in *A. thaliana* (Hirota et al., 2007; Chandler and Werr, 2017). *A. thaliana DREB2A* overexpression enhanced drought and heat tolerance in transgenic plants (Sakuma et al., 2006a,b). Additionally, DREB2A affected leaf senescence by interacting with radical-induced cell death 1 (RCD1) under heat stress (Vainonen et al., 2012). Thus, an ERF protein is often involved in several regulatory networks, which causes some ERF proteins to exert the same or opposite function during different processes. AtERF1 activates the defense-related gene, *PDF1.2* (Berrocal-Lobo et al., 2002), while AtERF4 represses *PDF1.2* in biotic stress tolerance (McGrath et al., 2005). Moreover, AtERF2 and AtERF5 are activators and AtERF3 is a repressor and they regulate downstream target genes during transcription in defense responses (Fujimoto et al., 2000).

Although different ERFs may exhibit opposing functions in different vital processes, an ERF may function in several of these processes. For example, *AtDREB1A* in transgenic *A. thaliana* plants resulted in the dwarfed phenotypes and freezing and dehydration tolerance, whereas *AtDREB2A* transgenic plants exhibited slight growth retardation (Liu et al., 1998). Additionally, wild-type *A. thaliana* plants overexpressing *AtERF53* exhibited unstable drought tolerance, while *rglg1rglg2* double mutant plants overexpressing *AtERF53* exhibited stable drought tolerance, as RGLG1 and RGLG2 together negatively regulate *AtERF53* transcription (Cheng et al., 2012). However, *AtERF7* overexpression decreased the sensitivity of guard cells to ABA and increased water loss during transpiration, which reduced drought tolerance in transgenic plants. Contrasting results were found in *AtERF7* RNA interference plants (Song et al., 2005). These studies suggest that although the regulatory pathways of DREB and ERF proteins differ, they can achieve the same effects in different vital processes.

Solanum lycopersicum (tomato), as an important fruit vegetable, is widely planted in many countries. Tomato fruit is highly nutritious and has a unique flavor, and can be eaten raw, boiled, or processed into ketchup or juice. Thus, improving the fruit yield and quality of tomato is the primary goal of tomato production. To achieve this goal, our understanding of the underlying molecular mechanisms of different vital processes must be enhanced, including seed germination, fruit ripening and softening, flower development, and defense responses to biotic and abiotic stresses. Among these processes, ERFs as regulators or repressors play important roles that affect different gene networks. In this study, we identified, corrected, and analyzed all ERF and DREB subfamily members based on *S. lycopersicum* genome database versions 2.0, 3.2, and 4.0. To understand the functions of ERF and DREB subfamily members, several RNA sequencing databases from NCBI SRA data library were used to analyze gene expression levels during the tomato growth and development. In addition, the inhibitory function of the EAR motif in several ERF proteins was tested by yeast one-hybrid assay, and gene expression profiles were analyzed by qRT-PCR during fruit ripening. These results will help establish the regulatory networks of ERF and DREB subfamilies, and uncover effective ways to improve tomato yield and quality.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The tomato (*S. lycopersicum*) 'AC' cultivar was grown in a greenhouse for 22 ∼ 25°C 16 h light (150 $\mu Em^{-2}s^{-1}$)/16 ∼ 18°C 8 h dark. Three samples of the green pulp and color-breaking pulp for 1, 3, 6, and 9 days were collected and utilized for gene expression analysis. The fresh young leaves, flowers, or shoot tips of tomato plants were gathered, frozen in liquid nitrogen, and utilized to clone the CDS of target genes.

Identification of ERF and DREB Subfamily Members in Tomato

The genome sequences of *S. lycopersicum* were downloaded from a database of gene annotations, SGN, (Release v2.0,

v3.2, and v4.0[1]). The hidden Markov model (HMM) profile of the AP2 domain (PF00847) was downloaded from the Pfam database[2]. HMMER v3.3 was used to search for candidate AP2/ERF genes from the tomato genome database. The default parameters were used and the cutoff value was set to 0.001. All of the candidate AP2/ERF proteins with only a single AP2 domain were selected as candidate ERF proteins. The Pfam, SMART[3], and NCBI CDD databases[4] were used to validate the candidate ERF proteins. Finally, the identification results of the 3 genome versions (2.0, 3.2, and 4.0) and NCBI database were compared to determine the final ERF subfamily members of *S. lycopersicum*.

Phylogenetic Analysis

Multiple sequence alignments of the tomato ERF proteins were performed using CLUSTAL W based on the complete sequences. To understand the relationship among the tomato ERF proteins, a phylogenetic tree was inferred using the maximum likelihood method based on the Whelan and Goldman model (Whelan and Goldman, 2001) of MEGA v7.0 with the following parameters: JTT + G model, partial deletion with 80% site coverage cutoff, and 1,000 bootstrap replications (Kumar et al., 2016).

Gene Structure and Conserved Motif Analyses

According to the cluster analysis results of the tomato ERF gene subfamily, the structural domain analysis of the ERF protein sequences of different groups was conducted using Jalview software (Waterhouse et al., 2009). Homologous alignments were compared using T-Coffee software (Notredame et al., 2000). The protein sequences of non-conservative regions were deleted. The alignment results were preserved in EPS format. Conserved motifs of the tomato ERF subfamily proteins were identified using the Multiple Em for Motif Elicitation (MEME) online tool v5.1.1[5] with the following parameters: number of occurrences of a single motif distributed among the sequences within the model, 0 or 1 per sequence; maximum number of motifs, 20; optimum width of each motif, 6–50 residues.

Transcriptome Data Source and Bioinformation Analysis

Transcriptome sequencing data were downloaded from the NCBI SRA database[6] using the SRA toolkit. The project number of transcriptome data used in this article is as follows: PRJNA507622 (*S. lycopersicum*, 30 tomato organs) (Penin et al., 2019) and PRJNA528656 (*S. lycopersicum*, fruit). Every organ includes three biological repetitions in these data.

The transcripts per million (TPM) expression values of the transcriptomes of different organs in tomato were obtained using the SRA toolkit and Salmon software (Patro et al., 2017). Subsequently, the TPM values were processed to quantify of gene expression levels of the original data. The expression heat map of the ERF genes in different organs was drawn using R-pheatmap based on the TPM values.

RNA Isolation and cDNA Synthesis

Total RNA was extracted using the TaKaRa MiniBEST Universal RNA Extraction Kit (TaKaRa, Kyoto, Japan). First-strand cDNA was synthesized using the PrimeScript^{TM} IV 1st strand cDNA Synthesis Mix (TaKaRa, Kyoto, Japan). The first-strand cDNA was utilized in the expression and amplification of *SlERF* genes.

Quantitative Real-Time PCR Assay

qRT-PCR was used to analyze the expression of 14 *SlERF* genes during tomato fruit ripening. cDNA was used as a template with the primer pairs shown in **Supplementary Table 1**. The *Sl-Actin* gene was used as internal control. The reactions were performed in triplicate for each sample using the TB Green® Premix Ex Taq^{TM} (Tli RNaseH Plus), Bulk (TaKaRa, Kyoto, Japan) on an qRT-PCR system under the following conditions: 95°C for 30 s, 40 cycles of 95°C for 5 s and 60°C for 30 s, and 1 cycle of 95°C for 15 s, 60°C for 1 min, 95°C for 15 s, and 60°C for 15 s. The dissociation curve was used to validate the specificity of each primer pair. Each experiment was repeated three times. The relative expression level of each gene was calculated and each result was reported as mean (±SE) of three independent experiments. ANOVA was used to identify statistically significant difference among genes ($P < 0.05$).

Yeast One-Hybrid Assay

To explore whether EAR motif-containing SlERF proteins played a repression role, 14 EAR motif-containing *SlERF* genes were selected to construct Y1H vectors by Matchmaker Gold Yeast One-Hybrid System Kit (Clontech, Mountain View, CA, United States). First, the pAbAi vector was cut by *Sal*I restriction enzyme, 3× DRE and 3× GCC elements were, respectively, inserted into the linear pAbAi vector, and then the recombinant vectors were transformed into Y1H gold yeast competent cells (primers in **Supplementary Table 1**). The yeast cells were selected on a plate without uracil for selective glucose synthesis and the positive yeast colonies with 3× DRE or 3× GCC elements were identified by the colony PCR analysis (Matchmaker Insert Check PCR Mix 1, Clontech, Mountain View, CA, United States). The yeast strains with 3× DRE or 3× GCC elements were cultured on SD/-Ura medium with 50, 100, and 150 ng/mL of aureobasidin A (AbA) to select the minimum inhibitory concentration. Second, the complete and deletion EAR motif of 14 SlERF CDS were amplified (primers in the **Supplementary Table 1**) and inserted into the SmaI-linearized pGADT7-Rec AD vector by the In-Fusion PCR Cloning Kits, and then the AD-prey recombinant vectors were transformed into the YIH gold yeast competent cells with 3× DRE or 3× GCC elements. The yeast cells were selected on an SD/-Leu/AbA plate and identified by the colony PCR analysis (Matchmaker Insert Check PCR Mix 2, Clontech, Mountain View, CA, United States). Each screening was performed three times.

[1]http://solgenomics.net/organism/solanum_lycopersicum/genome
[2]http://pfam.xfam.org/
[3]http://smart.embl-heidelberg.de/
[4]https://www.ncbi.nlm.nih.gov/cdd
[5]http://meme-suite.org/tools/meme
[6]https://www.ncbi.nlm.nih.gov/sra/

RESULTS

Sequence Correction of ERF and DREB Subfamily Genes

To ensure the sequence accuracy of all AP2/ERF genes, the Pfam model (pf00847) of the AP2 domain downloaded from the Pfam website was used to search the tomato v4.0 protein database. A total of 166 AP2/ERF proteins with an AP2 domain *E*-value < 0.001 were obtained. Among these proteins, 20 had ≥2 AP2 domains, while 146 proteins had single AP2 domain. Among the latter, 3 proteins with the B3 domain were RAV-type AP2/ERF proteins. Thus, there were 143 ERF subfamily proteins with a single AP2 domain. The 143 protein sequences were submitted to the Pfam, CDD, and smart websites for conservative domain analysis. Subsequently, 140 tomato ERF subunit proteins with a single AP2 domain were identified. The sequences of these proteins were compared in 3 tomato genome sequencing protein databases (versions 2.0, 3.2, and 4.1) (**Supplementary Table 2**); 27 genes were found to be different. The protein and CDS sequences of these 27 genes were compared and confirmed according to the tomato genome and NCBI databases (**Supplementary Tables 3, 4**). Finally, the corrected protein sequences were used for subsequent analyses.

Characteristics, Polarity, and Chemical Structure Analysis of the 14th and 19th Amino Acids in the AP2 Domain

Among the 140 ERF proteins with a single AP2 domain, the 14th amino acid of the AP2 domain was Valine (V) in 57 genes. Among these 57 proteins, the 19th amino acid of the AP2 domain was glutamic acid (E) in 30 proteins, aspartic acid (D) in 4 proteins, asparagine (N) in 1 protein, glutamine (Q) in 4 proteins, histidine (H) in 6 proteins, leucine (L) in 10 proteins, alanine (A) in 1 protein, and V in 1 protein (**Table 1** and **Supplementary Table 5**). These 57 proteins were identified as DREB proteins. Additionally, the 14th and 19th amino acids of the AP2 domain were isoleucine (I) and D, respectively, in SlERF2-5, SlERF10-6, and SlERF10-8. The codon of I was AUA/AUC, GUA/GUG/GUU/GUC for V, but GCA/GCG/GCU/GCC for A. The characteristics, polarity, and chemical structure of I and V were hydrophobic, non-polar, and aliphatic, while A was neutral, non-polar, and aliphatic (**Table 1** and **Supplementary Table 5**). Thus, I can only be a V mutation. Accordingly, the 3 proteins were identified as DREB proteins. In the 19th amino acid of the AP2 domain, the hydrophilic amino acids included E, D, N, Q, and H, the hydrophobic amino acids included L and V, and the neutral amino acids included A. The negative charged amino acids (E and D), uncharged amino acids (N and Q), and positively charged amino acids (H) were polar; the non-polar amino acids included L, A, and V. Additionally, H had a heterocycle chemical structure, while the others were aliphatic (**Table 1**). These differences may affect the functionality of DREB protein interactions with DRE and GCC boxes.

Among the 80 ERF subfamily members, the 14th and 19th amino acids of the AP2 domain were A and D in 70 proteins. Additionally, there was an A and tyrosine (Y) in 1 gene, A and N in 1 protein, threonine (T) and D in 1 protein, serine (S) and D in 4 proteins, E and D in 1 protein, glycine (G) and N in 2 proteins, and I and V in 1 protein (**Table 2** and **Supplementary Table 6**). In the 14th amino acid of the AP2 domain, the neutral amino acids included A, T, S, and G, and the hydrophilic amino acid included E. The non-polar amino acid was A, the polar amino acids without charges were T, S, and G, and the chemical structure of these amino acids is aliphatic. In the 19th amino acid of the AP2 domain, D, Y, and N comprised the hydrophilic amino acids, the negatively charged amino acid (D), the uncharged amino acids (Y and N) were polar, and the chemical structure of these amino acids was aliphatic (**Table 2**). Thus, the 80 proteins with a single AP2 domain were identified as ERF subfamily members. These differences may affect the functionality of ERF protein interactions with GCC boxes.

Phylogenetic Analysis of ERF and DREB Proteins

To understand their genetic relationships, the protein sequences of the 60 DREB and 80 ERF subfamily members were classified into 6 groups (**Figure 1**). The I group included 51 DREB proteins. Among these proteins, 37 and 14 proteins differentiated into the I-A and I-B subgroups, respectively. The I-A subgroup included

TABLE 1 | The 14th/19th amino acid analysis of the DREB subfamily AP2 domain.

Gene number	14th	19th	14th codon	19th codon	14th/19th Characters	14th polarity	14th chemical structure	19th polarity	19th chemical structure
30	V	E	GUA/GUG/GUU/GUC	GAA/GAG	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity with negative charge	Aliphatic
4	V	D	GUU	GAU/GAC	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity with negative charge	Aliphatic
1	V	N	GUU	AAC	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity, without charge	Aliphatic
4	V	Q	GUA/GUG/GUU	CAA	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity without charge	Aliphatic
6	V	H	GUU/GUC	CAU/CAC	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity with positive charge	Heterocycle
10	V	L	GUA/GUG/GUU/GUC	UUG/CUU/CUA/UUA	Hydrophobic/hydrophobic	Non-polarity	Aliphatic	Non-polarity	Aliphatic
1	V	A	GUA	GCA	Hydrophobic/neutral	Non-polarity	Aliphatic	Non-polarity	Aliphatic
1	V	V	GUG	GUU	Hydrophobic/hydrophobic	Non-polarity	Aliphatic	Non-polarity	Aliphatic
2	I	D	AUA	GAC	Hydrophobic/hydrophilic	Non-polarity	Aliphatic	Polarity with negative charge	Aliphatic
1	I	V	AUC	GUU	Hydrophobic/hydrophobic	Non-polarity	Aliphatic	Non-polarity	Aliphatic

TABLE 2 | The 14th/19th amino acid analysis of the ERF subfamily AP2 domain.

Gene number	14th	19th	14th codon	19th codon	Characters	14th polarity	14th chemical structure	19th polarity	19th chemical structure
70	A	D	GCA/GCG/GCU/GCC	GAU/GAC	Neutral/hydrophilic	Non-polarity	Aliphatic	Polarity with negative charge	Aliphatic
1	A	Y	GCA	UAU	Neutral/hydrophilic	Non-polarity	Aliphatic	Polarity without charge	Aromatic
1	A	N	GCU	AAU	Neutral/hydrophilic	Non-polarity	Aliphatic	Polarity without charge	Aliphatic
1	T	D	ACG	GAU	Neutral/hydrophilic	Polarity without charge	Aliphatic	Polarity with negative charge	Aliphatic
4	S	D	UCU/UCA	GAU/GAC	Neutral/hydrophilic	Polarity without charge	Aliphatic	Polarity with negative charge	Aliphatic
1	E	D	GAA	GAU	Hydrophilic/hydrophilic	Polarity with negative charge	Aliphatic	Polarity with negative charge	Aliphatic
2	G	N	GGA	AAC	Neutral/hydrophilic	Polarity without charge	Aliphatic	Polarity without charge	Aliphatic

30 proteins with V14E19, 3 proteins with V14Q19 (SlERF12-9, SlERF1-13, and SlERF7-1), 1 protein with V14A19 (SlERF11-4), 1 protein with V14V19 (SlERF1-5), and 2 proteins with V14L19 (SlERF6-5 and SlERF12-3) (**Supplementary Table 5**). Seven CBF proteins (SlERF3-7, SlERF3-22, SlERF3-6, SlERF8-2, SlERF8-3, SlERF12-11, and SlERF1-3) clustered together and were in the I-A subgroup (**Figure 1**). Additionally, the I-A subgroup included 4 proteins (SlERF9-1, SlERF2-10, SlERF4-10, and SlERF4-11) with EAR motif (DLNxxP or LxLxL) (**Supplementary Table 5**). However, the I-B subgroup only included 6 proteins with V14H19 and 8 proteins with V14L19. SlERF9-10 and SlERF8-14 showed EAR motif in the I-B subgroup (**Figure 1** and **Supplementary Table 5**).

Group II included 49 ERF subfamily proteins. Among these proteins, 14 and 35 proteins differentiated into the II-A and II-B subgroups, respectively (**Figure 1**). All members of the II-A subgroup belonged to ERF proteins with A14D19. In the II-B subgroup, there were 33 proteins with A14D19, 1 protein with T14D19 (SlERF1-10), and 1 protein with S14D19 (SlERF1-11) (**Supplementary Table 6**). In the II-B subgroup, 6 proteins with the EDLL transactivation motif (ExxxxDxxxLxxxL) clustered together (SlERF3-1, SlERF9-7, SlERF9-3, SlERF9-4, SlERF3-2, and SlERF9-8) (**Figure 1**). SlERF5-8 was also an EAR motif-containing protein in the II-B subgroup. Groups III and IV included 13 and 16 ERF subfamily proteins, respectively. However, group III also included a DREB protein with V14D19 (SlERF1-4) that clustered with an ERF-type protein (SlERF1-2). One protein with S14D19 (SlERF1-15) and 1 EAR motif-containing protein (SlERF4-1) were clustered into group III. Among these proteins in group IV, there were 2 proteins with S14D19 (SlERF3-16 and SlERF12-1), 2 proteins with G14N19 (SlERF12-6 and SlERF12-7), and all others belonged to proteins with A14D19. Additionally, 9 EAR motif-containing proteins (SlERF10-1, SlERF7-5, SlERF12-1, SlERF7-2, SlERF2-6, SlERF7-3, SlERF10-2, SlERF3-4, and SlERF3-16) were in group IV (**Figure 1** and **Supplementary Table 6**). Group V had 8 DREB proteins, including 3 proteins with V14D19, 1 protein with V14Q19, 1 protein with V14N19, 2 proteins with I14D19, and 1 protein with I14V19 (**Figure 1** and **Supplementary Table 5**). Group VI included 1 protein with E14D19 (SlERF2-1) and 1 protein with A14D19 (SlERF9-2) (**Supplementary Table 6**).

Motif Analysis of ERF and DREB Protein Sequences

To understand the constructional characteristics of ERF and DREB proteins, a Multiple Em for Motif Elicitation (MEME) analysis was conducted to calculate the possible motifs of the 140 proteins. β1 of the AP2 domain was located on the left of motif 2, β2 was located on the right of motif 2 and left of motif 3, and β3 and α were located on motif 1. All 140 proteins, except SlERF9-1, SlERF10-9, SlERF6-1, SlERF8-4, SlERF8-12, SlERF3-8, SlERF2-1, and SlERF9-2, had motifs 1, 2, and 3. SlERF9-1, SlERF10-9, SlERF6-1, SlERF8-4, and SlERF8-12 had only motifs 1 and 3, as well as a same sequence to motif 25 in front of motif 3. SlERF3-8 had motifs 2, 3, and 16. Motif 16 had a similar sequence as motif 1. SlERF2-1 and SlERF9-2 did not have motifs 1, 2 and 3, but had a similar sequence as motif 16 with motif 1 (**Figure 2**). These results suggest that SlERF9-1, SlERF10-9, SlERF6-1, SlERF8-4, SlERF8-12, SlERF3-8, SlERF2-1, and SlERF9-2 may have the low ability to bind with GCC or DRE boxes.

In addition to motifs 1, 2, 3, 16, and 25, some motifs were located on both sides of the AP2 domain of many ERF and DREB proteins. For example, motifs 10 and 20 were near the left of the AP2 domain in 7 CBF and SlERF1-13, respectively, while motifs 4 and 5 were near the right of motif 1 and especially motif 4 (**Figure 2**). These findings suggest that motifs 4, 5, 10, and 20 may be involved in the process of AP2 domain binding with GCC or DRE boxes. However, motifs 6–9, 11–15, 17–19, 21, 23, and 24 were relatively far away from the AP2 domain, may be located in the transactivation or repression domains, and may be involved in regulating the expression of their downstream target genes. However, some ERF and DREB proteins, including SlERF4-10, SlERF7-1, SlERF10-5, SlERF6-7, SlERF3-15, SlERF6-9, SlERF3-16, SlERF12-6, SlERF12-7, SlERF3-4, SlERF7-3, SlERF2-6, SlERF3-13, and SlERF2-5, did not have other motifs, except motifs 1, 2, 3, and 4 (**Figure 2**). Nevertheless, a few of these proteins had a typical EAR motif, including SlERF4-10, SlERF3-16, SlERF3-4, SlERF7-3, and

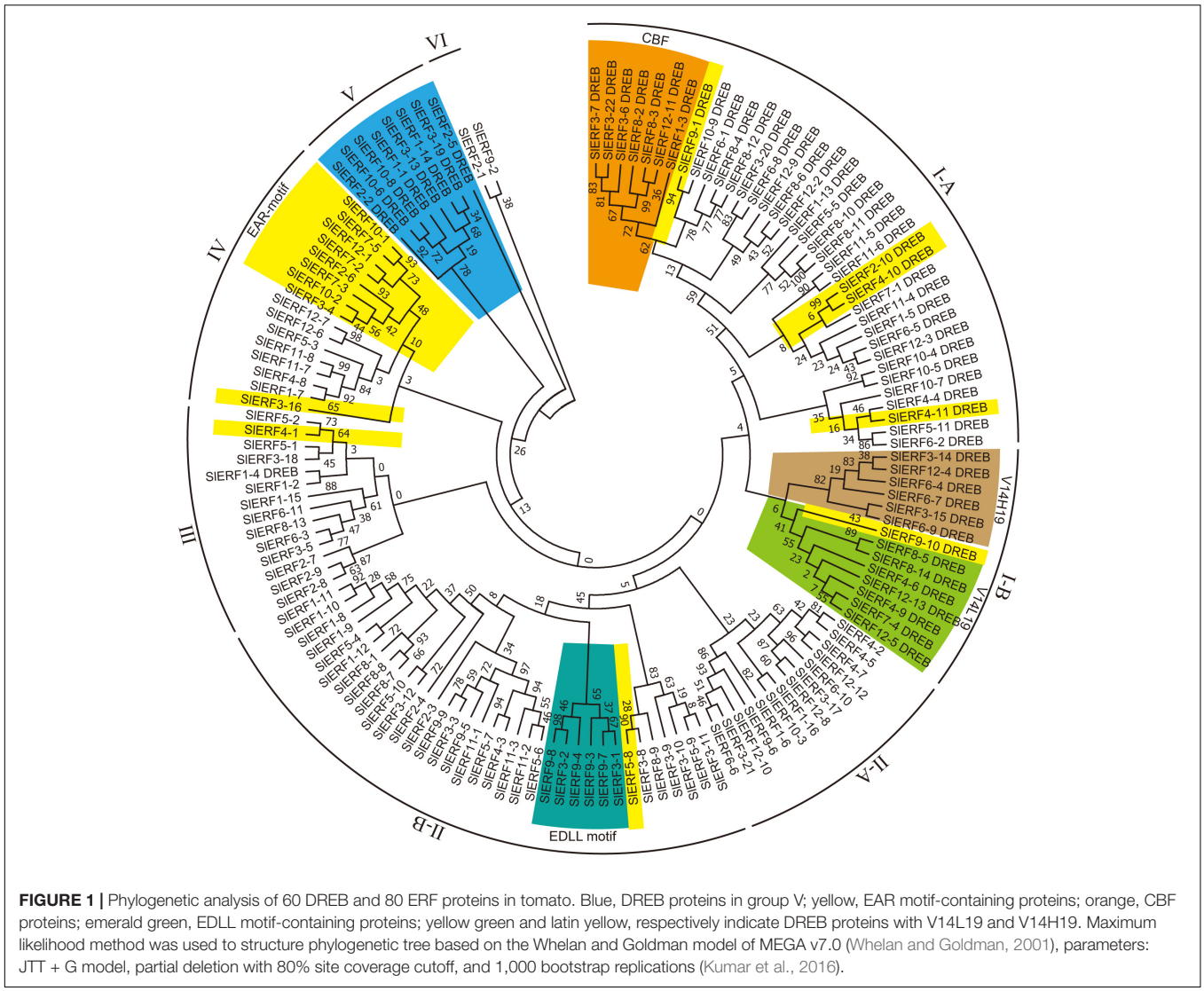


FIGURE 1 | Phylogenetic analysis of 60 DREB and 80 ERF proteins in tomato. Blue, DREB proteins in group V; yellow, EAR motif-containing proteins; orange, CBF proteins; emerald green, EDLL motif-containing proteins; yellow green and latin yellow, respectively indicate DREB proteins with V14L19 and V14H19. Maximum likelihood method was used to structure phylogenetic tree based on the Whelan and Goldman model of MEGA v7.0 (Whelan and Goldman, 2001), parameters: JTT + G model, partial deletion with 80% site coverage cutoff, and 1,000 bootstrap replications (Kumar et al., 2016).

SlERF2-6. These proteins bound to DNA with the AP2 domain and repressed the expression of downstream target genes with the EAR domain. However, SlERF7-1, SlERF10-5, SlERF6-7, SlERF3-15, SlERF6-9, SlERF12-6, SlERF12-7, SlERF3-13, and SlERF2-5 especially protein sequences with <100 amino acids (SlERF12-6 and SlERF12-7) may competitively inhibit other ERF and DREB proteins (**Figure 2**).

The Special Amino Acid of AP2 Domain May Affect Protein Binding With DRE/CRT and GCC Boxes

Previous studies found that some DREB and ERF subfamily proteins only bound to DRE or GCC boxes, but most of these proteins can also interact with these boxes. However, the correlation between the characteristics and binding ability of DREB and/or ERF subfamily proteins remains unclear. To distinguish the difference between DREB and ERF proteins during binding with DRE/CRT or GCC boxes, the AP2 domain amino acid sequences of 49 *A. thaliana* and 19 other species ERF proteins, including 8 tomato ERF proteins, were compared. The binding assays of the 68 ERF proteins with DRE and GCC boxes were completed through an electrophoretic mobility shift assay (EMSA), yeast one-hybrid, or proteome chip assays. Among these proteins, there were 42 protein AP2 domains that included P9, 5 included H9, 5 included S9, 6 included N9, 3 included Q9, 2 included K9, 2 included T9, and 1 included I9 (**Figure 3**). Only 19 proteins bound with GCC box, including 17 ERF with P9, 2 DREB with 1 P9 and 1 H9. Additionally, 37 proteins bound with DRE/CRT and GCC boxes, including 23 ERF with P9, 14 DREB with 1 P9, 4 H9, 4 N9, 2 Q9, 1 T9, 1 K9, and 1 I9. Only 12 proteins bound with DRE, including 1 ERF with P9, 11 DREB with 5 S9, 2 N9, 2 K9, 1 Q9, and 1 T9 (**Figure 3**). These results suggest that almost all ERFs with P9 and H9 can interact with GCC box, and most can also bind with DRE/CRT. All DREB with S9 can interact with DRE/CRT, but other DREBs with N9, K9, Q9, T9, and I9 can also bind with DRE/CRT and/or GCC boxes. The A14 and A15 amino acids of ERF AP2 domain were conserved, but the

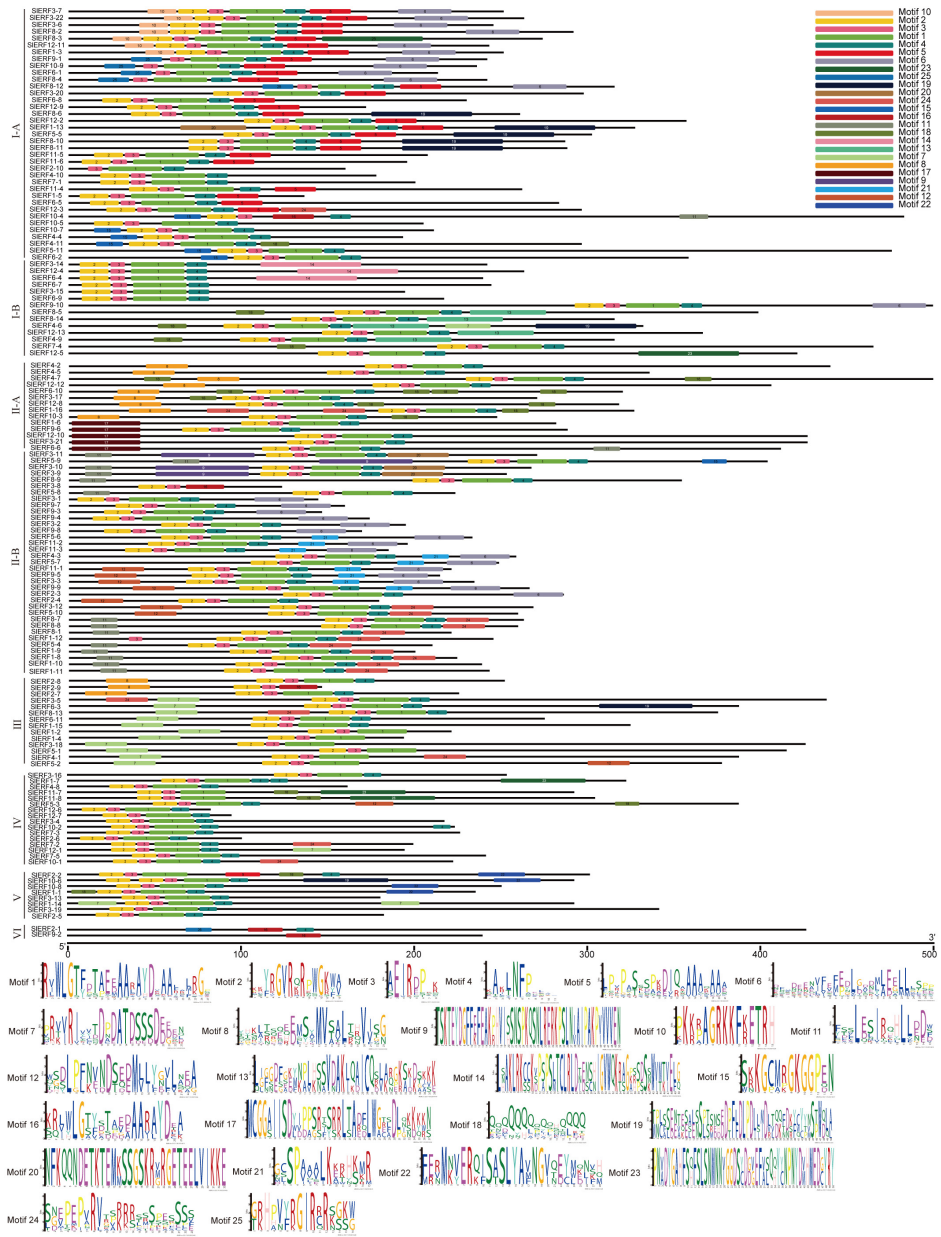


FIGURE 2 | Twenty-five motif predictive analysis of 60 DREB and 80 ERF proteins in tomato using a Multiple Em for Motif Elicitation (MEME).

13th amino acid may be Y, F, or W. The W13 and V14 amino acids of the DREB AP2 domain were conserved, but the 15th amino acid may be S, A, or C (**Figure 3**). These characteristics of ERFs and DREBs may affect the ability of proteins to bind with DRE/CRT and GCC boxes.

In tomato DREB subfamily members, there are 10 DREBs with S9W13V14S15 (SlERF3-20, SlERF5-5, SlERF6-8, SlERF8-10, SlERF8-11, SlERF8-12, SlERF9-1, SlERF10-9, SlERF11-4, and SlERF12-9), 1 DREB with S9W13V14C15 (SlERF8-4) (**Supplementary Table 3**), and 1 DREB with S9W13I14A15 (SlERF10-6), which suggests that these 11 DREBs may bind with DRE/CRT. There were 4 DREBs with H9W13V14S15 (SlERF12-4, SlERF6-7, SlERF3-15, and SlERF6-9) and 9 DREB with H9W13V14A15 (SlERF3-14, SlERF9-10, SlERF8-5, SlERF8-14, SlERF4-6, SlERF12-13, SlERF4-9, SlERF7-4, and SlERF12-5) (**Supplementary Table 3**), which suggest that these interact with the GCC box. Seven CBF proteins exist in tomato DREB subfamily members, including 5 CBF with N9W13V14C15 (SlERF3-7, SlERF3-22, SlERF3-6, SlERF8-2, and SlERF1-3) and 2 CBF with D9W13V14C15 (SlERF8-3 and SlERF12-11). There were 3 DREB with N9W13V14S15 (SlERF6-1, SlERF8-6, and SlERF1-13), 5 DREB with K9W13V14S15 (SlERF11-5, SlERF11-6, SlERF1-5, SlERF6-5, and SlERF12-3), 4 DREB with K9W13V14A15 (SlERF2-10, SlERF4-10, SlERF7-1, and

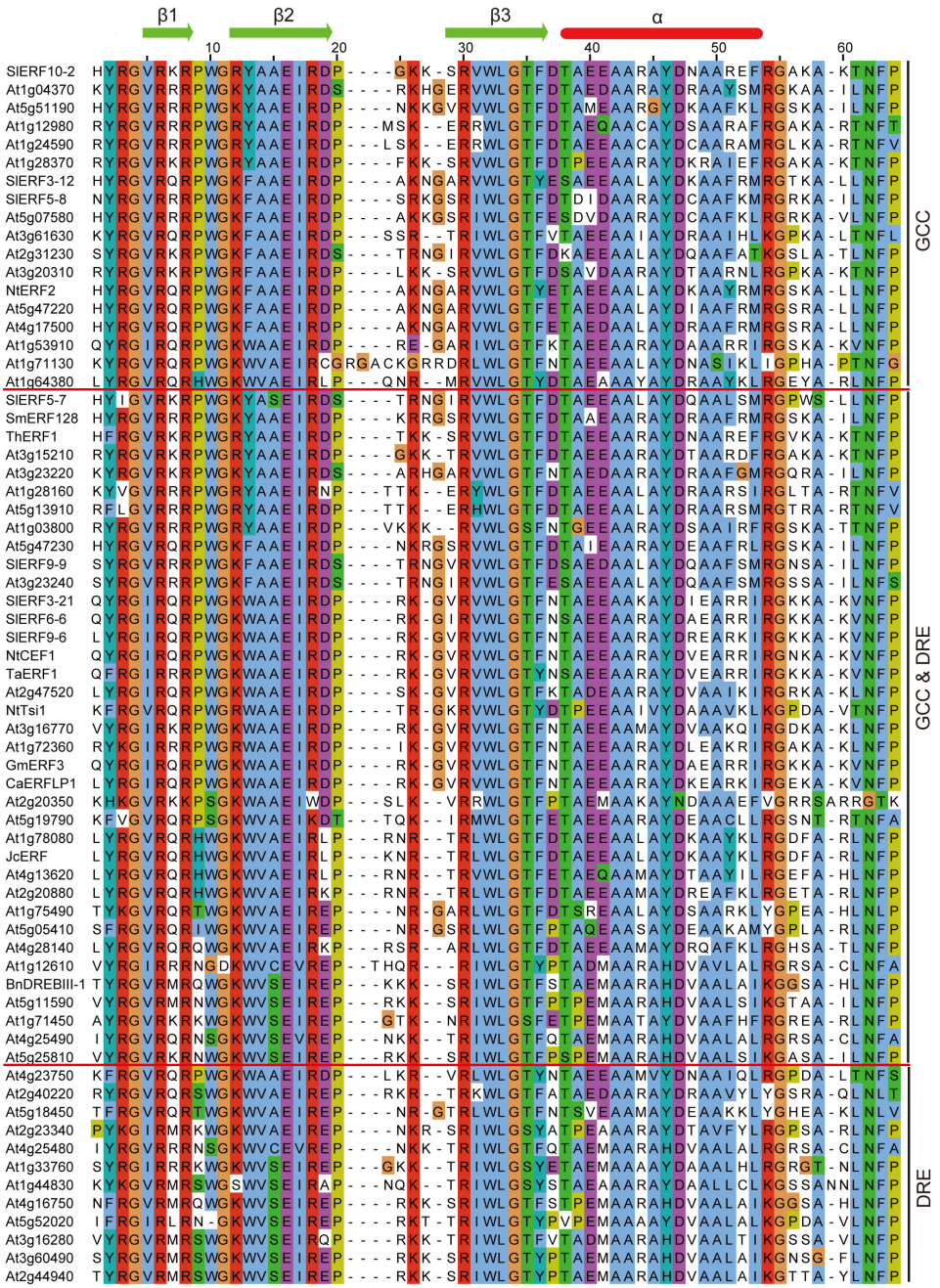


FIGURE 3 | Comparison of the AP2 domain amino acid sequences of 49 *A. thaliana* and 19 other species ERF proteins. SlERF10-2, SlERF3-12, and SlERF5-8 (Tournier et al., 2003); At1g04370, At5g51190, At1g12980, At1g24590, At1g28370, At5g07580, At3g61630, At2g31230, At3g20310, At4g17500, At1g53910, At1g71130, At1g64380, At3g23220, At1g28160, At5g13910, At1g03800, At2g20350, At5g19790, At4g13620, At2g20880, At1g75490, At4g28140, At1g12610, At1g71450, At4g25490, At2g40220, At5g18450, At2g23340, At4g25480, At1g33760, At1g44830, At4g16750, At5g52020, At3g16280, At3g60490, At4g23750, and At2g44940 (Gong et al., 2008); NtERF2 (Solano et al., 1998); At5g47220, At3g15210, and At5g47230 (Fujimoto et al., 2000); SlERF5-7 (Huang et al., 2004), SmERF128 (Zhang et al., 2019), ThERF1 (Wang et al., 2014), SlERF9-9 (Klay et al., 2018), At3g23240 (Solano et al., 1998), SlERF3-21 (Wu et al., 2008), SlERF6-6 (Zhang et al., 2004), SlERF9-6 (Zhang Z. et al., 2009), NtCEF1 (Lee et al., 2005), TaERF1 (Xu et al., 2007), At2g47520 (Lee et al., 2015), NtTsi1 (Park et al., 2001), At3g16770 (Büttner and Singh, 1997), At1g72360 (Licausi et al., 2010), GmERF3 (Zhang G. et al., 2009), CaERFLP1 (Jae-Hoon et al., 2004), At1g78080 (Lin et al., 2008), JcERF (Tang et al., 2007), At5g05410 (Sakuma et al., 2002), BnDREBIII-1 (Liu et al., 2006), At5g11590 (Wei et al., 2005), and At5g25810 (Sun et al., 2008).

SlERF1-4), 6 DREB with T9W13V14A15 (SlERF10-4, SlERF10-5, SlERF10-7, SlERF4-4, SlERF4-11, and SlERF5-11), 3 DREB with 1 Q9W13V14S15 (SlERF6-4), 1 I9W13V14A15 (SlERF6-2), and 1 A9W13V14S15 (SlERF12-2), 5 DREB with P9W13V14A15 (SlERF2-2, SlERF1-1, SlERF3-13, SlERF1-14, and SlERF3-19), and 1 DREB with P9W13I14A15 (SlERF10-8) (**Supplementary**

Table 3). These DREB proteins may interact with DRE/CRT, some of which may also bind with GCC box.

In tomato ERF subfamily members, there were ERFs with 22 P9Y13A14A15, 21 P9F13A14A15, and 21 P9W13A14A15 (**Supplementary Table 3**). There were ERFs with K9Y13A14 A15 (SlERF3-18), P9Y13A14S15 (SlERF5-7), P9Y13G14A15 (SlERF12-7), P9Y13G14V15 (SlERF12-6), T9F13A14A15 (SlERF3-2), K9F13A14T15 (SlERF5-1), Q9F13S14A15 (SlERF1-11), P9F13S14A15 (SlERF12-1), Q9F13T14A15 (SlERF1-10), E9W13A14A15 (SlERF1-2), 2 with K9W13A14A15 (SlERF4-1 and SlERF5-2), and 2 with P9W13S14A15 (SlERF1-15 and SlERF3-16) (**Supplementary Table 3**). This indicates that P9A14A15 of the ERF AP2 domain may play an important role in binding with GCC box.

Approximately 21 ERF and 15 DREB Genes Are Regulated by Auxin and/or Ethylene During Tomato Fruit Development and Ripening

To explore the relationship of ERF family members with fruit development and ripening, the tomato fruit RNA-Seq data from NCBI SRA were analyzed (**Supplementary Table 7**). Among the 140 *SlERF* genes, the expression of 36 genes showed some regular changes from 8 to 53 days post anthesis (DPA). In the I-A group, the transcripts per million (TPM) values of *SlERF4-10*, *SlERF10-4*, *SlERF5-11*, and *SlERF6-2* increased during fruit development (from 8 to 28 DPA) and during ripening (from 28 to 53 DPA), whereas that of *SlERF8-2* and *SlERF11-4* did not change during fruit development and increased during ripening (**Figure 4A** and **Supplementary Table 7**). In the 5,000-bp promoter sequence of these genes, *SlERF8-2*, *SlERF11-4*, *SlERF10-4*, and *SlERF6-2* had 8, 2, 1, and 5 ARF binding sites, respectively; *SlERF4-10* had 6 ARF, 5 DRE/CRT, and 1 GCC box binding sites; *SlERF5-11* had 2 ARF and 1 DRE/CRT binding sites (**Figure 4B**). In addition, *SlERF4-10* had an EAR motif. These results suggest that *SlERF4-10* and *SlERF5-11* could be directly induced by IAA signal during fruit development and ET signal during ripening; *SlERF8-2* and *SlERF11-4* just responded to the ET signal during ripening, and *SlERF10-4* and *SlERF6-2* could be directly induced by IAA signal during development and indirectly regulated by ET interaction with other hormone signals during ripening (**Figure 4C**). In addition, the TPM values of *SlERF6-8*, *SlERF12-9*, and *SlERF5-5* showed the opposite trend with the above six genes, and those of *SlERF12-2* and *SlERF1-13* decreased during fruit development and ripening (**Figure 4A** and **Supplementary Table 7**). *SlERF6-8* and *SlERF12-9,* respectively, had two and three ARF binding sites; *SlERF5-5* had six ARF and one DRE/CRT binding sites; *SlERF1-13* had one DRE/CRT binding site; *SlERF12-2* had three ARF and one DRE/CRT binding sites (**Figure 4B**). These results indicated that IAA signal directly promoted *SlERF6-8*, *SlERF12-9*, and *SlERF5-5* and inhibited *SlERF12-2* and *SlERF1-13* expression, but ET signal directly repressed *SlERF5-5*, *SlERF12-2*, and *SlERF1-13* expression and indirectly inhibited *SlERF6-8* and *SlERF12-9* expression (**Figure 4C**).

In the I-B group, the TPM values of *SlERF12-4* and *SlERF6-4* slightly increased during fruit development and markedly reduced during fruit ripening (**Figure 4A** and **Supplementary Table 7**). Moreover, the *SlERF12-4* promoter showed four ARF, two DRE/CRTs, and one GCC box binding sites (**Figure 4B**). Interestingly, one GCC box closed to one ARF binding sites between −4,491 and −4,501 bp, and one ARF site showed a 5-bp overlap with one DRE/CRT sites within −190 to −197 bp (**Supplementary Tables 8, 9**). The *SlERF6-4* promoter showed four ARF and two DRE/CRT binding sites (**Figure 4B**), and two ARF site showed a 5-bp overlap with two DRE/CRT sites within −259 to −252 bp and −136 to −129 bp (**Supplementary Tables 8, 9**). These results suggested that *SlERF12-4* and *SlERF6-4* were directly induced by IAA signal and inhibited by ET signal (**Figure 4C**). The TPM value of *SlERF8-14* did not clearly show regularity during fruit development, but significantly increased during fruit ripening (**Figure 4A** and **Supplementary Table 7**). However, its promoter only showed three ARF and no DRE/CRT or GCC box binding sites (**Figure 4B**), which indicated that *SlERF8-14* may be indirectly induced by ET signal interacting with other hormones during ripening and was not affected by IAA signal during fruit development.

In the II-A group, the TPM values of *SlERF4-7* and *SlERF3-17* did not visibly change during fruit development but significantly increased during fruit ripening; the TPM values of *SlERF12-12*, *SlERF1-6*, and *SlERF6-6* continuously increased during fruit development and ripening; the TPM value of *SlERF1-6* first increased and then decreased during fruit development, and continuously increased during fruit ripening (**Figure 4A**). In addition, *SlERF4-7* promoter had seven ARF and two DRE/CRT binding sites. Moreover, one ARF site showed a 5-bp overlap with one DRE/CRT site from −1,623 to −1,629 bp (**Supplementary Tables 8, 9**). The *SlERF12-12* promoter had six ARF, two DRE/CRT, and one GCC box binding sites, and one ARF site showed a 5-bp overlap with one DRE/CRT site from −438 to −444 bp (**Supplementary Tables 8, 9**). The *SlERF1-6* promoter had one ARF and one DRE/CRT binding sites. In addition, the promoters of *SlERF3-17* and *SlERF6-6* had nine and two ARF binding sites, respectively (**Figure 4B**). These results indicate that ET signal may directly induce *SlERF4-7*, *SlERF12-12*, and *SlERF1-6* expression and indirectly regulate *SlERF3-17*, and *SlERF6-6*; IAA signal directly improved *SlERF12-12*, *SlERF6-6*, and *SlERF1-6* expression but did not affect *SlERF4-7* and *SlERF3-17* expression during fruit development (**Figure 4C**).

In the II-B group, the TPM values of *SlERF5-9*, *SlERF5-8*, and *SlERF2-4* almost showed a continuously increasing trend from fruit development to ripening, while that of *SlERF5-10* and *SlERF8-8* did not change during fruit development and obviously increased during ripening (**Figure 4A**). Moreover, the *SlERF5-9* promoter had two ARF and one DRE/CRT binding sites, and one ARF site showed a 5-bp overlap with one DRE/CRT site from −3,318 to −3,324 bp (**Supplementary Tables 8, 9**). The promoters of *SlERF5-8*, *SlERF2-4*, *SlERF5-10*, and *SlERF8-8* had 5, 7, 0, and 1 ARF binding sites, respectively (**Figure 4B**). Four ARF sites in the *SlERF2-4* promoter were detected between −4,244 and −4,363 bp. In addition, *SlERF5-8* had an EAR motif. These results suggested that IAA signal could directly induce *SlERF5-9*, *SlERF5-8*, and *SlERF2-4* expression and did not affect *SlERF5-10* and *SlERF8-8* expression during fruit

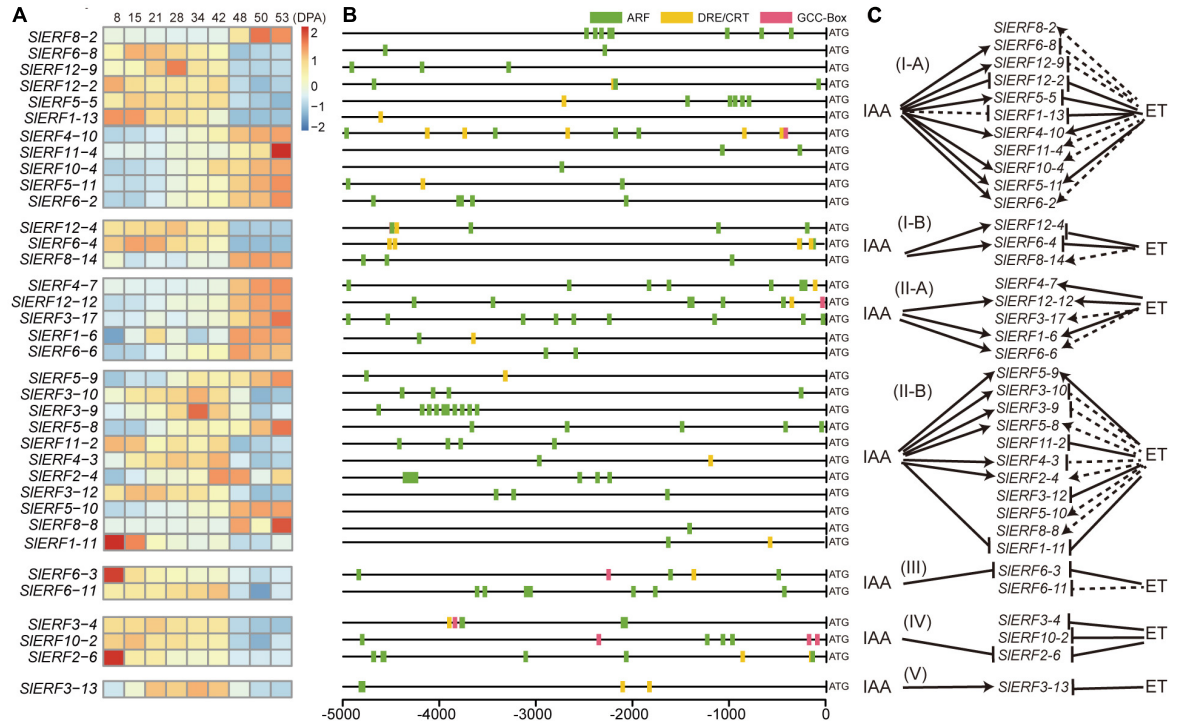


FIGURE 4 | The relationship of DREB and ERF subfamily genes with tomato fruit development and ripening. **(A)** Heatmap analysis of 36 *SlERF* genes based on TPM values from 8 to 53 DPA. **(B)** ARF, DRE/CRT, and GCC boxes analysis of 36 *SlERF* gene promoters. **(C)** Relationship of 36 *SlERF* genes with IAA and ET signals. Black solid arrow, indicates direct positive regulation; black dashed arrow, indicates indirect positive regulation; black solid lines ending with bar, indicates direct repression; black dashed lines ending with bar, indicates indirect repression.

development; ET signal directly induced *SlERF5-9* transcription and indirectly enhanced *SlERF5-8*, *SlERF2-4*, *SlERF5-10*, and *SlERF8-8* expression by interacting with other hormones during fruit ripening (**Figure 4C**). However, the TPM values of *SlERF3-10*, *SlERF3-9*, and *SlERF4-3* visibly increased during fruit development and decreased during fruit ripening; the TPM values of *SlERF11-2* and *SlERF3-12* did not change during development and decreased during ripening (**Figure 4A** and **Supplementary Table 7**). Among these five genes, the promoters of *SlERF3-10*, *SlERF3-9*, *SlERF11-2*, and *SlERF3-12* had 4, 10, 4, and 3 ARF binding sites, respectively. The *SlERF4-3* promoter had one ARF and one DRE/CRT binding sites (**Figure 4B**). In addition, 9 ARF sites of the *SlERF3-9* promoter were clustered between −3,611 and −4,185 bp. The TPM value of *SlERF1-11* continuously decreased from fruit development to ripening. The *SlERF1-11* promoter had one ARF and one DRE/CRT binding sites. These results suggested that *SlERF3-10* and *SlERF3-9* could be directly induced by IAA signal during fruit development and indirectly inhibited by ET signal during fruit ripening; *SlERF11-2* and *SlERF3-12* expressions were not affected by IAA signal and were indirectly inhibited by ET signal; *SlERF1-11* expression was directly inhibited by IAA and ET signals, while *SlERF4-3* expression was directly induced by IAA during fruit development and repressed by ET during fruit ripening (**Figure 4C**).

In the III group, the TPM value of *SlERF6-3* was reduced during fruit development and ripening (**Figure 4A** and **Supplementary Table 7**). Moreover, the *SlERF6-3* promoter had four ARF, two DRE/CRT, and one GCC box binding sites (**Figure 4B**), and one ARF site showed a 5-bp overlap with one DRE/CRT site at −1,607 to −1,613 bp (**Supplementary Tables 8, 9**). This indicated that *SlERF6-3* may be negatively regulated by IAA and ET. In addition, the TPM value of *SlERF6-11* did not show obvious changes during fruit development and reduced during fruit ripening; the *SlERF6-11* promoter had seven ARF binding sites (**Figures 4A,B**). These findings suggested that IAA signal do not regulate *SlERF6-11* expression through the ARF pathway, but ET signal may indirectly inhibit *SlERF6-11* expression via other pathways (**Figure 4C**).

In the IV group, the TPM values of *SlERF3-4* and *SlERF10-2* did not obviously change during fruit development and decreased during fruit ripening (**Figure 4A**). The *SlERF3-4* promoter had four ARF, two DRE/CRT, and two GCC box binding sites, and one ARF, one DRE/CRT, and one GCC box sites were situated within the −3,760 to −3,776 bp (**Supplementary Tables 8, 9**). The *SlERF10-2* promoter had four ARF, one DRE/CRT, and three GCC box binding sites, and one ARF site showed a 5-bp overlap with one DRE/CRT site within −1,065 to −1,071 bp (**Supplementary Tables 8, 9**). These results suggested that *SlERF3-4* and *SlERF10-2* were not regulated by IAA signal during development but directly repressed by ET signal during ripening (**Figure 4C**). In addition, the TPM value of *SlERF2-6* continuously decreased from fruit development to

ripening (**Figure 4A** and **Supplementary Table 7**). The *SlERF2-6* promoter had seven ARF and two DRE/CRT binding sites, and four ARF sites were located within −4,574 to −4,689 bp (**Supplementary Tables 8, 9**). These results indicate that *SlERF2-6* expression could be directly inhibited by IAA signal during fruit development and by ET signal during ripening (**Figure 4C**).

In the V group, the TPM value of *SlERF3-13* (DREB) increased during fruit development and decreased during ripening (**Figure 4A** and **Supplementary Table 7**). Moreover, its promoter had two ARF and two DRE/CRT binding sites (**Figure 4B**). These findings suggested that *SlERF3-13* could be directly induced by IAA signal during fruit development and directly repressed by ET signal during ripening (**Figure 4C**).

Expression Analysis of EAR Motif-Containing *SlERF* Genes in Tomato

Among the 140 SlERF proteins, 11 ERF and 5 DREB subfamily proteins that have one or two typical EAR (LxLxL or DLNxxP) motifs were found. SlERF10-1, SlERF7-5, SlERF12-1, SlERF7-2, SlERF4-11, SlERF9-1, SlERF2-10, and SlERF4-10 have a DLNxxP motif in C-terminal; SlERF2-6, SlERF3-16, and SlERF4-1 have an LxLxL motif in C-terminal, but SlERF9-10 and SlERF5-8 have an LxLxL motif in N-terminal. In addition, SlERF7-3, SlERF10-2, and SlERF3-4 show not only an independent LxLxL motif but also a DLNxxP motif in C-terminal. However, the DLNxxP motif of SlERF7-3 and SlERF10-2 connect with an LxLxL sequences and form a strong repressive motif (**Figure 5A**).

To understand the function of the 16 EAR motif-containing *SlERF* genes, the TPM values of these genes were analyzed in flowers, fruits, meristems, seeds, leaves, and roots (**Supplementary Table 10**). Among these genes, *SlERF10-1*, *SlERF7-5*, *SlERF12-1*, *SlERF7-2*, *SlERF7-3*, *SlERF10-2*, and *SlERF2-10* were largely expressed in almost all organs; *SlERF2-6*, *SlERF3-4*, *SlERF4-10*, and *SlERF5-8* were expressed in most organs such as flowers, seeds, and leaves, but *SlERF2-6* and *SlERF3-4* were not expressed in fruits, in contrast to *SlERF4-10* and *SlERF5-8*. In addition, *SlERF3-16* was only expressed in floral and vegetative meristems, seeds, and young leaves; *SlERF4-11* was only expressed in seeds, *SlERF8-14* and *SlERF4-1* were low expressed in almost all organs, and *SlERF9-10* and *SlERF9-1* were not expressed in almost all organs (**Figure 5B** and **Supplementary Table 10**). These results indicate that *SlERF10-1*, *SlERF7-5*, *SlERF12-1*, *SlERF7-2*, *SlERF7-3*, *SlERF10-2*, and *SlERF2-10* are involved in regulating the development of almost all organs, other genes except *SlERF9-10* and *SlERF9-1* play roles in regulating the development of part organs.

Among the 16 EAR motif-containing *SlERF* genes, at least 12 genes were expressed in every organ, especially 15 genes in green large seeds (**Figure 5C** and **Supplementary Table 10**). In addition, most genes without the EAR motif were expressed in every organ such as the maximum 95 genes in mature petals and the minimum 66 in red pulp (**Figure 5C** and **Supplementary Table 10**). However, total TPM values of 16 ERFs with the EAR motif showed a very high ratio of all ERF genes in every organ, and exceeded genes without the EAR motif in most organs such as 55.84% in mature flowers, 57.53% in floral meristem, 56.50% in vegetative meristem, and 58.72% in cotyledons (**Figure 5D** and **Supplementary Table 10**). These results suggested that the EAR motif-containing *SlERF* genes may play important roles in balancing regulatory function of other ERF and DREB subfamily genes to downstream target genes during tomato growth and development.

The EAR Motif of Most SlERFs Functions in Gene Repression in Tomato

To understand whether the EAR motif plays a repressing role in ERF proteins regulating the expression of their target genes, a yeast one-hybrid experiment (Y1H) was used in this study. Out of 16 ERF proteins with the EAR motif, 14 these were selected for the construction of the *pGADT-SlERF* and *pGADT-SlERF-N* carriers. In the *pGADT-SlERF-N* carriers, the EAR sequences of the ERF proteins were deleted (**Figure 6A**). In Y1H, the yeast cells with *pBait-AbAi-3* × *DRE* and *pGADT* carrier did not grow in the yeast medium supplemented with 100 ng/mL AbA. The yeast cells with *pBait-AbAi-3* × *DRE* together with *pGADT-SlERF2-6*, *pGADT-SlERF2-10*, *pGADT-SlERF3-16*, *pGADT-SlERF4-1*, *pGADT-SlERF4-10*, *pGADT-SlERF4-11*, *pGADT-SlERF5-8*, *pGADT-SlERF7-3*, or *pGADT-SlERF7-5* carriers also showed the same appearance with the yeast cells including *pBait-AbAi-3* × *DRE* and *pGADT* carrier (**Figure 6B**). However, after the EAR sequences of these proteins were deleted, parts of these grew in the yeast medium supplemented with 100 ng/mL AbA such as SlERF2-6, SlERF4-1, SlERF4-10, SlERF4-11, and SlERF7-3. In addition, the number of the yeast cells with *pBait-AbAi-3* × *DRE* together *pGADT-SlERF3-4* or *pGADT-SlERF10-1* was low, but that of yeast cells without the EAR sequences of the two proteins significantly increased in the yeast medium supplemented with 100 ng/mL AbA. The yeast cells with *pGADT-SlERF7-2*, *pGADT-SlERF7-2-N*, *pGADT-SlERF12-1*, or *pGADT-SlERF12-1-N* carriers showed a few colonies, and those of *pGADT-SlERF2-10*, *pGADT-SlERF2-10-N*, *pGADT-SlERF3-16*, *pGADT-SlERF3-16-N*, *pGADT-SlERF5-8*, *pGADT-SlERF5-8-N*, *pGADT-SlERF7-5*, or *pGADT-SlERF7-5-N* did not grow in the yeast medium supplemented with 100 ng/mL AbA (**Figure 6B**). These results suggested that the EAR motifs of the SlERF2-6, SlERF4-1, SlERF4-10, SlERF4-11, SlERF7-3, SlERF3-4, and SlERF10-1 proteins repressed the expression of their target genes, of which the promoters had DRE/CRT elements, with those of SlERF7-2 and SlERF12-1 imparting weak effects, and the DNA binding domain of SlERF2-10, SlERF3-16, SlERF5-8, and SlERF7-5 did not interact with the DRE element.

In the yeast cells with the *pBait-AbAi-3* × *GCC* and *pGADT* carrier, AbA could inhibit cells growth. This appearance also showed in the yeast cells with *pBait-AbAi-3* × *GCC* together *pGADT-SlERF3-4*, *pGADT-SlERF4-11*, *pGADT-SlERF7-2*, *pGADT-SlERF7-3*, *pGADT-SlERF7-5*, or *pGADT-SlERF10-2*. However, after the EAR sequences of these genes were deleted, the yeast cells including SlERF4-11-N, SlERF7-2-N, or SlERF7-3-N proteins could grow; those of SlERF3-4-N, SlERF7-5-N, and SlERF10-2-N proteins did not grow in the yeast medium supplemented with 100 ng/mL AbA (**Figure 6B**). The

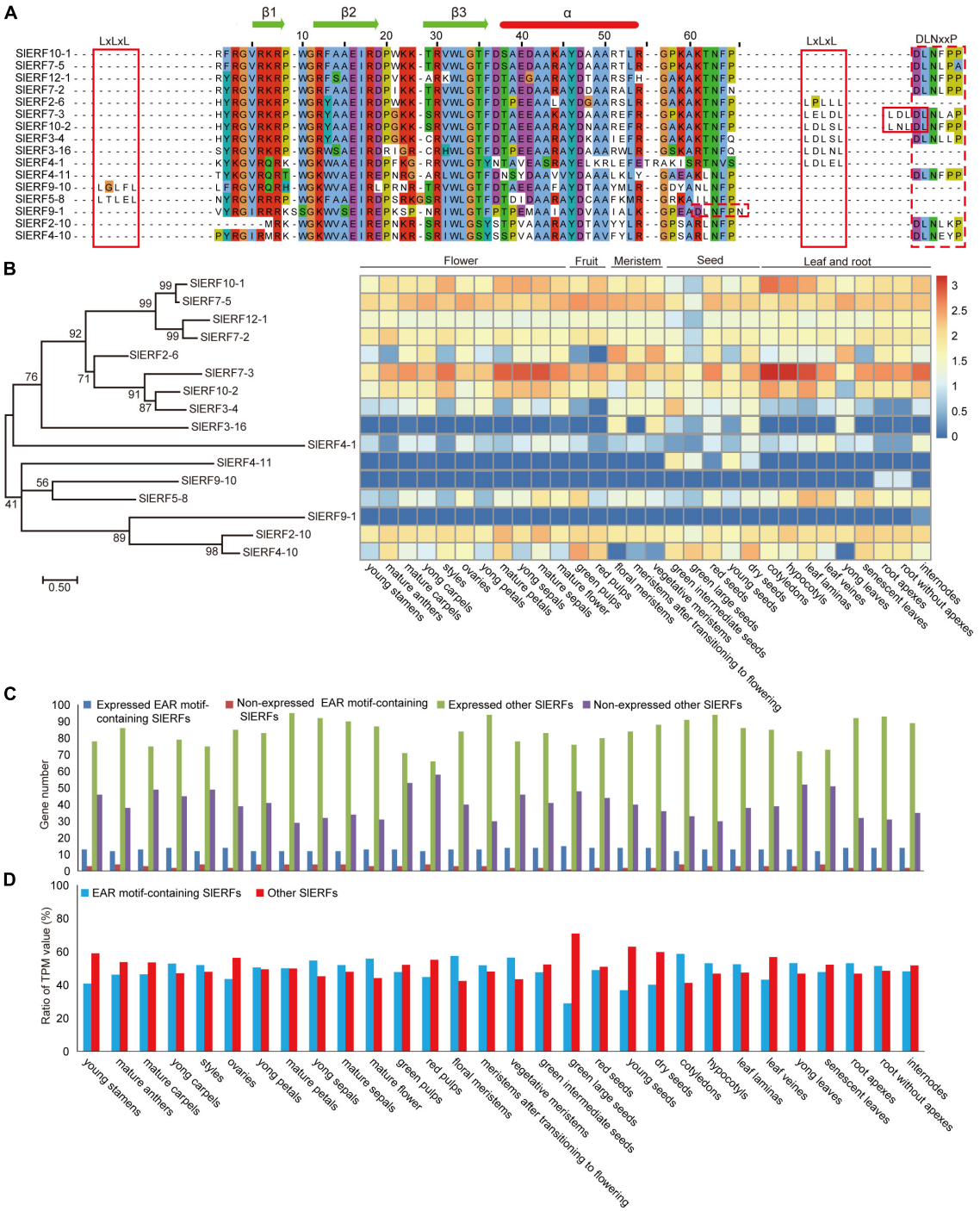


FIGURE 5 | Expression levels of the EAR motif-containing *SlERF* genes in 30 tomato organs. **(A)** AP2 domain and EAR motifs analysis in the EAR motif-containing SlERF proteins. **(B)** Heatmap analysis of all EAR motif-containing *SlERF* genes based on TPM values in 30 tomato organs. **(C)** Gene numbers analysis of the expressed and non-expressed genes among EAR motif-containing *SlERFs* and other *SlERF* genes in 30 tomato organs. **(D)** The total TPM ratio analysis of the EAR motif-containing *SlERFs* and other *SlERF* genes in 30 tomato organs.

yeast cells with *pBait-AbAi-3* × *GCC* together *pGADT-SlERF2-6*, *pGADT-SlERF2-10*, *pGADT-SlERF3-16*, *pGADT-SlERF4-1*, *pGADT-SlERF4-10*, *pGADT-SlERF5-8*, *pGADT-SlERF10-1*, or *pGADT-SlERF12-1* carriers could grow in the yeast medium supplemented with 100 ng/mL AbA. After the EAR sequences of these genes were deleted, the yeast cells, including SlERF2-6-N, SlERF3-16-N, SlERF4-1-N, or SlERF12-1-N proteins, showed more growth, but those of SlERF2-10-N, SlERF5-8-N, or

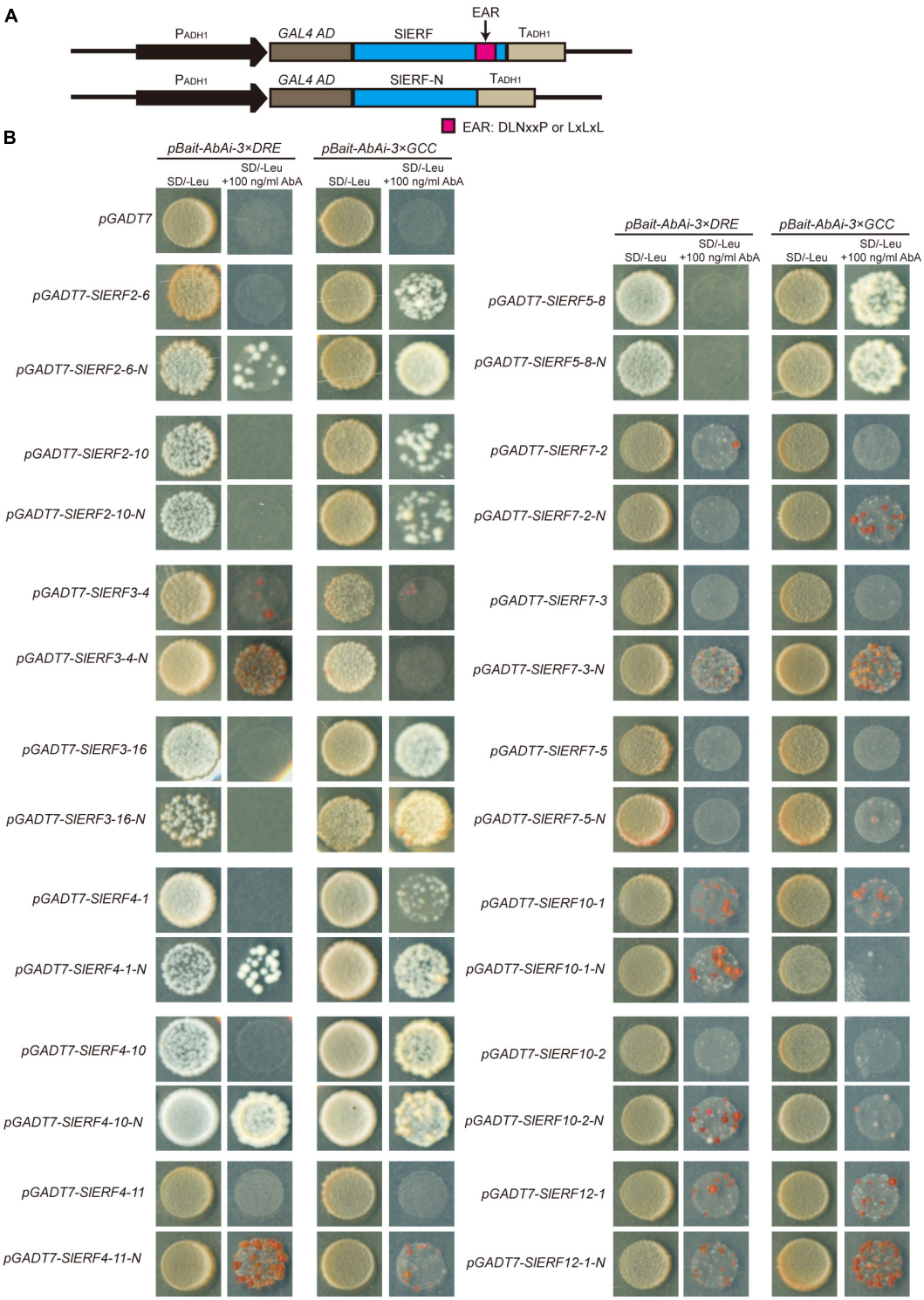


FIGURE 6 | The repressing function analysis of the EAR motif in SlERF proteins by a Y1H assay. **(A)** Construction of the Y1H recombinant vectors. **(B)** The activation ability analysis of 14 SlERF proteins retained or deleted EAR motifs in a Y1H assay.

SlERF4-10-N proteins showed scarce changes compared to SlERF2-10, SlERF5-8, or SlERF4-10 proteins. In addition, yeast cells, including SlERF10-1-N proteins, did not even grow in the yeast medium supplemented with 100 ng/mL AbA (**Figure 6B**). These results indicated that the EAR motifs of SlERF2-6, SlERF3-16, SlERF4-1, SlERF4-11, SlERF7-2, SlERF7-3,

and SlERF12-1 proteins repressed target gene promoters within the GCC box, and those of SlERF2-10, SlERF4-10, and SlERF5-8 might not effectively offset the activation of the AD domain when ERF proteins bind to the GCC box, and the AP2 domain of SlERF3-4, SlERF7-5, SlERF10-1, and SlERF10-2 might not or weakly bind to the GCC box.

Some EAR Motif-Containing SlERFs Are Involved in Fruit Ripening

To investigate whether ERFs with the EAR motif regulated fruit ripening, the expression levels of 14 ERFs with the EAR motif were analyzed by qRT-PCR at fruit different ripening stages (**Figure 7A**). The results indicated that the expression levels of *SlERF2-10*, *SlERF7-2*, and *SlERF12-1* did not change (**Figures 7B,C**). In addition, *SlERF9-1* and *SlERF9-10* were not expressed in tomato fruits (**Figure 7B**). Therefore, these five genes were not involved in regulating fruit repining. At the beginning of the breaker stage (BR), *SlERF4-10, SlERF7-3,* and *SlERF10-1* were up-regulated, whereas *SlERF2-6*, *SlERF3-4*, *SlERF3-16*, *SlERF4-1*, *SlERF4-11*, *SlERF5-8*, and *SlERF7-5* were down-regulated. Three days after the BR, only the expression levels of *SlERF4-10* and *SlERF7-3* increased, whereas that of other genes, except for *SlERF2-10*, *SlERF7-2*, *SlERF10-1*, and *SlERF12-1*, decreased. Six days after the BR, *SlERF4-10, SlERF5-8, SlERF7-3,* and *SlERF10-2* were up-regulated, and whereas the other 5 *SlERF* genes, except for *SlERF2-10*, *SlERF4-1*, *SlERF7-2*, *SlERF10-1*, and *SlERF12-1*, were down-regulated. Nine days after the BR, only *SlERF5-8* and *SlERF10-2* were up-regulated, *SlERF2-10*, *SlERF4-1*, *SlERF4-10*, *SlERF7-2*, *SlERF7-3*, and *SlERF10-1* did not change, and six other genes were down-regulated (**Figures 7D,E**). These results indicate that *SlERF4-10* and *SlERF7-3* played regulatory roles at the BR for 1–6 days, *SlERF5-8* and *SlERF10-2* at the BR for 6–9 days, *SlERF10-1* only at the BR for 3 days, *SlERF2-10*, *SlERF7-2*, and *SlERF12-1* were not affected with the increase in ethylene release, while *SlERF2-6*, *SlERF3-4*, *SlERF3-16*, *SlERF4-1*, *SlERF4-11*, and *SlERF7-5* were inhibited by ethylene.

DISCUSSION

The Ability of ERFs Binding With DNA Was Not Dependent on the AP2 Domain but Required the Motif Near the AP2 Domain

Ethylene responsive factor, as a specific transcription factor in plant, is involved in the life cycle of higher plants, including growth, development, abiotic, and biotic stresses (Chen Y. et al., 2021; Ji et al., 2021; Tian et al., 2021; Xing et al., 2021). However, the ERF family always includes numerous members in many plants (Zhu et al., 2018; Xu et al., 2020; Zong et al., 2021), while their functions are hard to distinguish in different life activities. Although the tomato genome has been sequenced and analyzed three times so far, there are still errors in the sequence analysis of many genes. In this study, we identified and analyzed 140 *SlERF* genes from tomato genome data v4.0 and then found that 26 of these showed errors in genome data v2.0, v3.2, or v4.0 such as initial codon missing prediction, part sequences deletion, intron missing prediction, and nucleotide deletion (**Supplementary Table 2**). Thus, correcting these errors as much as possible will help us better investigate *SlERF* functions.

In early reports, the 14th and 19th amino acids of AP2 domain were used to distinguish DREB and ERF subfamily proteins including V14E19, and A14D19 were, respectively, named DREB and ERF proteins. Moreover, DRE/CRT and GCC boxes only, respectively, interact with DREB and ERF proteins (Sakuma et al., 2002). For the 14th amino acid of AP2 domain, valine (V) had hydrophobic, non-polar, and aliphatic properties, while alanine (A) was neutral, non-polar, and aliphatic (**Tables 1**, **2**). For the 19th amino acid of AP2 domain, both glutamic acid (E) and aspartic acid (D) were hydrophilic, aliphatic, and shared polarity with negative charge (**Tables 1**, **2**). Therefore, the 14th amino acid of the AP2 domain could play an important role in ERF protein binding with DRE/CRT and GCC boxes, while the E19 or D19 of the AP2 domain might not affect this binding ability of ERF proteins with DRE/CRT and GCC boxes. For example, At5g19790 with V14D19 and At1g75490 with V14E19 could interact with both DRE/CRT and GCC boxes (Gong et al., 2008). Among 140 SlERF proteins, 60 and 80 were, respectively, divided into the DREB and ERF subfamilies. However, there were only 30 DREB proteins with V14E19 and 70 ERF with A14D19. In addition, the 19th amino acid of AP2 domain in some DREB proteins showed D, N, G, H, L, A, or V; that in some ERF proteins showed Y or N; the 14th amino acid of AP2 domain in two DREB proteins showed I and that in some ERF proteins showed T, S, E, and G. Property differences of these amino acids could have a certain effect when proteins bind with DRE/CRT and GCC boxes. In tomato, SlERF3-12 and SlERF5-8 with A14D19 only bind to the GCC box (Tournier et al., 2003), but SlERF9-9, SlERF3-21, SlERF6-6, and SlERF9-6 with A14D19 could interact with both DRE/CRT and GCC boxes (Zhang et al., 2004; Zhang Z. et al., 2009; Wu et al., 2008; Klay et al., 2018). In addition, SlERF5-7 with S14D19 was also demonstrated to bind with both DRE/CRT and GCC boxes (Huang et al., 2004). Furthermore, the AP2 domain consisted of three β-sheets and one α-helix; the arginine and tryptophan residues in the β-sheet could bind to the GCC box (Allen et al., 1998). In *Brassica napus*, BnDREBIII-1 interacted with DRE/CRT and GCC boxes, in contrast with BnDREBIII-4. The difference in their AP2 domain involved the 37th amino acid of the AP2 domain located in the α-helix (Liu et al., 2006). In this study, we compared the AP2 domain sequences of 68 ERF and DREB proteins from 20 species; the ability of these proteins binding with DRE/CRT and GCC boxes was assessed by EMSA, YIH, or proteome chip assays. We found that ERF proteins with P9A14A15 showed strong binding ability with the GCC box; in contrast, proteins with S/N/Q/K9V14S15 showed strong binding ability with DRE/CRT. However, some ERF proteins with P9A14A15 or S/N/Q/K9V14S15 could interact with DRE/CRT and GCC boxes (**Figure 3**). In addition, the MEME analysis results suggest that motifs 2, 3, 1, and 4 located in the AP2 domain, others are distributed outside the AP2 domain. Thus, motifs relatively distant from the AP2 domain such as motifs 7, 8, 9, 11, 12, 14, 17, 18, 19, 22, 23, and 24 may also act as a part

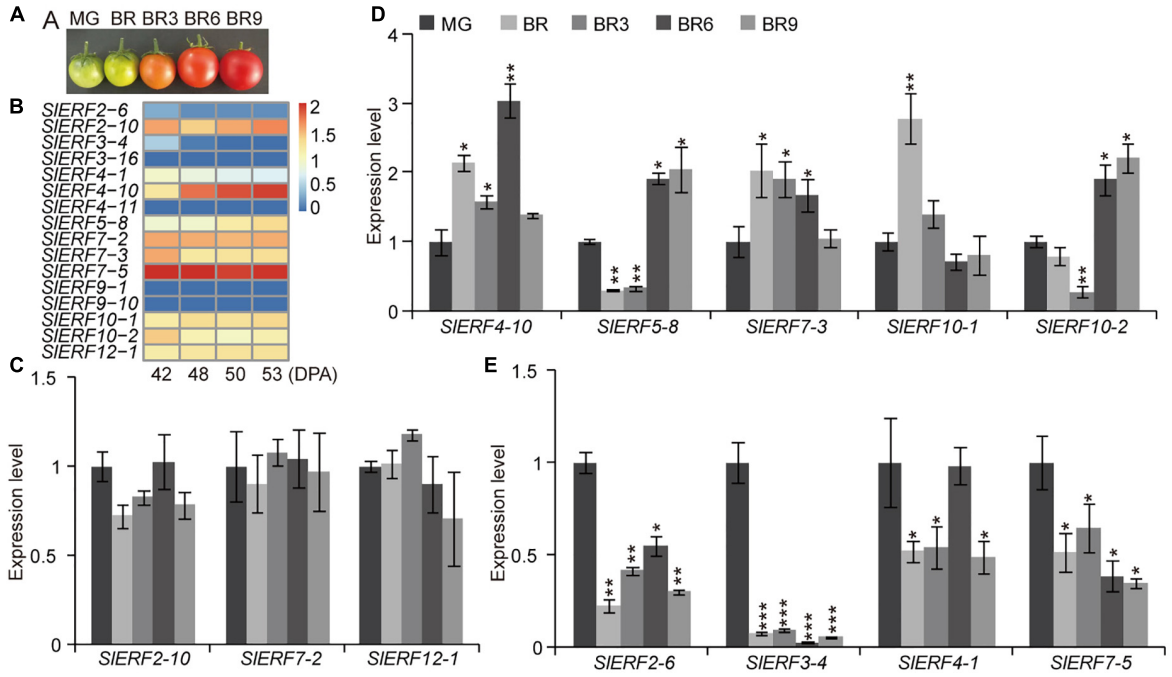


FIGURE 7 | Expression levels of the EAR motif-containing *SlERF* genes in tomato fruit ripening. **(A)** Tomato fruit ripening images. **(B)** Heatmap analysis of the EAR motif-containing *SlERF* genes based on TPM values in tomato fruit ripening. **(C–E)** The expression levels of 12 EAR motif-containing *SlERF* genes in tomato fruit ripening. MG, mature green; BR, the beginning of the breaker stage; BR3, BR6, and BR9, 3, 6, and 9 days after the BR, respectively. Student's *t*-test: $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$.

of transactivation domain; those near the AP2 domain such as motifs 5, 10, 13, 15, 20, 24, and 25 may help the AP2 domain to bind with DRE/CRT and GCC boxes (**Figure 2**). For example, the PKKPAGR signature sequence of *A. thaliana* CBF1 located left of the AP2 domain, mutations within this motif reduce CBF1 binding with DRE/CRT element and the expression level of COR gene (Canella et al., 2010). Thus, the ability of the ERF protein to bind to DNA was not dependent on the AP2 domain but involved other amino acids near the AP2 domain.

ERFs Are Involved in Fruit Development and Ripening Through Responses to Auxin and Ethylene Signals

The IAA promotes tomato fruit development, while the ET release improves fruit ripening. However, it is a complex regulatory mechanism that involves several metabolic networks. For example, MADS-box transcription factor ripening inhibitor (RIN) mutant showed incompletely ripened fruits (Ito et al., 2015, 2017), tomato organelle RNA recognition motif-containing proteins 4 (ORRM4) positively regulated fruit ripening (Yang et al., 2017), and DNA demethylase gene *SlDML2* was involved in DNA methylation of many genes during ripening (Lang et al., 2017). In this study, 36 ERF genes showed regular changes during tomato fruit development and ripening. The promoter of 35 of these showed ARF binding elements, but only 19 ERF genes were up-regulated during fruit development, which indicated that these genes may be directly induced by IAA signal (**Figures 4A,B**). However, *SlERF12-2*, *SlERF6-3*, and *SlERF2-6* were down-regulated during fruit development, suggesting that IAA signal negatively regulate their transcription. During fruit ripening, 17 ERF genes were up-regulated, but only promoters of *SlERF1-6*, *SlERF4-10*, *SlERF5-11*, *SlERF4-7*, *SlERF12-12*, and *SlERF5-9* showed the DRE/CRT and/or GCC boxes. As such, these may be directly induced by other ERF proteins. Interestingly, promoters of other 11 genes had no DRE/CRT and/or GCC boxes, indicating that these may be induced to regulate fruit ripening by ET interacting with other hormone pathways (**Figures 4A–C**). In addition, 19 ERF genes were inhibited by ET signal; 12 of these showed the DRE/CRT and/or GCC boxes and others did not harbor these elements (**Figures 4A,B**). These results suggested that ET signal directly or indirectly repressed these ERF gene expressions through other ERF proteins or other hormone pathways to improve fruit ripening. Early reports have demonstrated that *SlERF5-8* (*Sl-ERF.B3*) directly regulate *Sl-Aux/IAA27* to integrate ethylene and auxin signaling in tomato seedling development (Liu et al., 2018), while our results indicated that *SlERF5-8* expression continuously increased during fruit development and ripening, and its promoter had 5 ARF and no ERF binding elements (**Figures 4A,B**). In addition, the SlERF5-8 protein also had an EAR motif (**Figure 5A**). Thus, *SlERF5-8* and IAA and ET signals were found to have a negative feedback regulation relationship. Moreover, *SlERF5-10* (*SlPti4*)-silenced plants showed ABA accumulation, decrease in ET signals, and orange fruits; *SlERF5-10* could be induced by

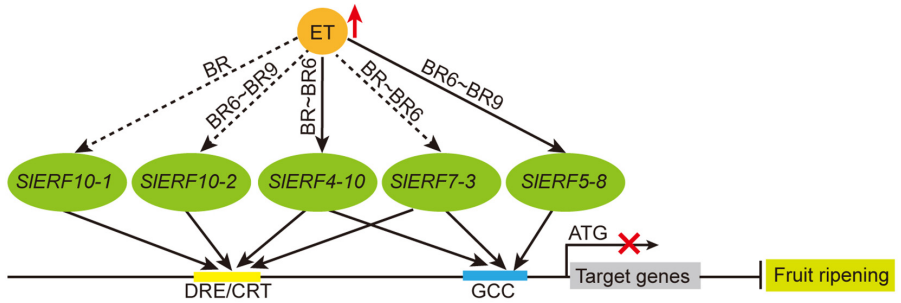


FIGURE 8 | The relationship of ET and EAR motif-containing SlERF genes during tomato fruit ripening. Red upward arrow, indicates increase; black solid arrow between ET and SlERF genes, indicates direct induction; black dashed arrow, indicates indirect induction; black solid arrow between SlERF and DRE/CRT or GCC boxes, indicates interaction; black line ending with bar, indicates repression; red cross, indicates transcriptional inhibition.

exogenous 1-aminocyclopropane 1-carboxylate (ACC) (Sun et al., 2018). However, *SlERF5-10* expression increased during ripening, but its promoter did not have ARF, DRE/CRT, and GCC boxes. These results indicated that *SlERF5-10* and ET signal had a feedback regulation relationship, i.e., *SlERF5-10* positively regulated ET biosynthesis and then ET signal indirectly induced *SlERF5-10* transcription during ripening. Thus, our results suggested that some ERF genes can be directly or indirectly induced by IAA and/or ET signals and feedback regulate IAA or ET biosynthesis during fruit development and ripening in tomato.

Most EAR Motif-Containing SlERFs as Inhibitors Balance Other ERF Functions During Tomato Growth and Development

The EAR motif-containing proteins repressed transcription of their target genes through recruiting histone deacetylase (Yang et al., 2018), but proteins without EAR motif did not have the same recruiting ability. In plants, the EAR motif-containing proteins play many important roles in regulating plant growth, development, and defense response (Chen et al., 2008; Jin et al., 2018; Lakehal et al., 2020; Lim et al., 2020). For example, *AtERF115* can repress adventitious rooting in *A. thaliana* through the JA and cytokinin signaling pathways (Gu et al., 2002); *AtERF4* inhibits the expression of JA-responsive defense gene and antagonizes JA inhibition of root elongation (McGrath et al., 2005). In this study, we found that only 11 ERF and 5 DREB proteins were present that included EAR motif in tomato (**Figures 1**, **5A**), but it is unclear whether these ERF proteins balanced with other ERFs without EAR motif proteins by competing between common target sites. In this study, we found that at least 13 EAR motif-containing ERF genes were expressed in several floral organs, red pulp, cotyledons, and senescent leaf, and at most 15 EAR motif-containing ERF genes were expressed in green large seeds (**Figure 5C**). However, most of other 124 ERF genes are expressed in every organ, such as the maximum 95 genes in mature petals and the minimum 66 genes in red pulp (**Figure 5C**). In addition, the TPM value ratio of EAR motif-containing ERF genes exceeded that of others in tomato 16 organs, especially in cotyledons (**Figure 5D**). Therefore, the EAR motif-containing ERF proteins may control the activation function of other ERF and DREB proteins by competition the binding sites of their target genes, and finally balance the expression of their target genes to ensure normal growth and development in tomato.

However, whether these proteins act as repressors remains unclear. Thus, 14 of these genes (*SlERF9-1* and *SlERF9-10* were hardly expressed in the tomato 30 tissues) were analyzed in a YIH experiment. Our results indicated that SlERF2-10, SlERF3-16, and SlERF5-8 only interacted with GCC box and its activity was hardly affected by EAR motif (**Figure 6B**). This may be because EAR motifs do not effectively offset the activity function of AD domain. However, the activities of other 10 proteins, except for SlERF7-5, were more or less affected by the EAR motif. SlERF3-4, SlERF4-11, SlERF7-3, SlERF7-5, and SlERF10-2 hardly promoted GAL4 gene expression, but they deleted EAR motif significantly enhanced GAL4 gene expression under 100 ng/mL AbA condition (**Figure 6B**). Besides, SlERF3-4 only interacted with DRE/CRT element, while SlERF7-3 could bind with DRE/CRT and GCC boxes (**Figure 6B**). Early studies indicated that the expression level of *SlERF10-2* (*LeERF3b*) was markedly increased in low-ethylene tomato fruit containing an ACC oxidase sense-suppression transgene as well as the ethylene insensitive mutant never ripe (Nr) (Ouyang et al., 2016). However, our results showed that *SlERF10-2* was firstly down-regulated in the BR 3 days and then up-regulated in the BR 6–9 days (**Figure 7B**). A similar trend was also found in the transcriptome data (**Figure 7A**). In addition, the overexpression of *SlERF10-1* (*SlERF36*) caused early flowering and plants senescence, and affected stomatal density, photosynthesis, and plant growth (Guo and Wang, 2011). In this study, the expression of *SlERF10-1* was only increased in the early BR and subsequently recovered the normal level during fruit ripening (**Figure 7B**). It suggested that *SlERF10-1* and *SlERF10-2* repressed the expression of genes related to fruit ripening. In addition, the expression of *SlERF4-10*, *SlERF5-8*, and *SlERF7-3* also increased during fruit ripening (**Figure 7B**). Moreover, *SlERF4-10* and *SlERF5-8* could be induced by ET signal (**Figures 4A–C**). Expressions of *SlERF2-10*, *SlERF7-2*, and *SlERF12-1* were not affected (**Figure 7C**), and those of *SlERF2-6*, *SlERF3-4*, *SlERF4-1*, and *SlERF7-5* were suppressed by ET

signal during fruit ripening (**Figure 7D**). Thus, the increase of ET directly promotes *SlERF4-10* and *SlERF5-8* and indirectly induces *SlERF7-3*, *SlERF10-1*, and *SlERF10-2* expression, these SlERF proteins inhibit their target gene transcriptions through binding with the DRE/CRT and/or GCC boxes of target gene promoters and then improve fruit ripening (**Figure 8**), while others did not regulate ripening or act in some basal metabolism.

Collectively, this work highlighted that there is scope to further understand the balance relationship of EAR motif-containing ERF and other ERF proteins. Our work showed the potential function of ERF and DREB subfamily genes to regulate fruit development and ripening in tomato and demonstrate the inhibition function of EAR motif in tomato ERF proteins. Our findings can further expand our understanding of the function of ERF and DREB subfamily genes in tomato.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

LZ, MQ, and YDL conceived and designed this study. LC, SP, QZ, SQ, and YFL carried out genome-wide identification and sequence analysis of the AP2/ERF genes. TX, LZ, and YDL carried out expression analysis of the AP2/ERF genes. LZ, MQ, and YDL drafted and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31772295, U1708232, 31991184, 32172554, and 32160713) and China Agriculture Research System (Grant No. CARS-25).

ACKNOWLEDGMENTS

We thank LetPub (www.letpub.com) for providing its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.849048/full#supplementary-material

REFERENCES

Akhter Most, S., Mohammed, N., Kouji, S., Takumi, S., Hiroaki, K., Takahide, S., et al. (2011). Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol.* 52, 344–360. doi: 10.1093/pcp/pcq196

Allen, M. D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M. (1998). A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J.* 17, 5484–5496. doi: 10.1093/emboj/17.18.5484

Berrocal-Lobo, M., Molina, A., and Solano, R. (2002). Constitutive expression of ETHYLENE-RESPONSE-FACTOR1 in Arabidopsis confers resistance to several necrotrophic fungi. *Plant J.* 29, 23–32. doi: 10.1046/j.1365-313x.2002.01191.x

Büttner, M., and Singh, K. B. (1997). Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 94, 5961–5966. doi: 10.1073/pnas.94.11.5961

Canella, D., Gilmour, S. J., Kuhn, L. A., and Thomashow, M. F. (2010). DNA binding by the Arabidopsis CBF1 transcription factor requires the PKKP/RAGRxKFxETRHP signature sequence. *Biochim. Biophys. Acta* 1799, 454–462. doi: 10.1016/j.bbagrm.2009.11.017

Chandler, J. W., and Werr, W. (2017). DORNRÖSCHEN, DORNRÖSCHEN-LIKE, and PUCHI redundantly control floral meristem identity and organ initiation in *Arabidopsis*. *J. Exp. Bot.* 68, 3457–3472. doi: 10.1093/jxb/erx208

Chen, G., Hu, Z., and Grierson, D. (2008). Differential regulation of tomato ethylene responsive factor LeERF3b, a putative repressor, and the activator Pti4 in ripening mutants and in response to environmental stresses. *J. Plant Physiol.* 165, 662–670. doi: 10.1016/j.jplph.2007.03.006

Chen, H. C., Chien, T. C., Chen, T. Y., Chiang, M. H., Lai, M. H., and Chang, M. C. (2021). Overexpression of a novel ERF-X-type transcription factor, OsERF106MZ, reduces shoot growth and tolerance to salinity stress in rice. *Rice* 14:82. doi: 10.1186/s12284-021-00525-5

Chen, Y., Shu, P., Wang, R., Du, X., Xie, Y., Du, K., et al. (2021). Ethylene response factor AcERF91 affects ascorbate metabolism via regulation of GDP-galactose phosphorylase encoding gene (AcGGP3) in kiwifruit. *Plant Sci.* 313:111063. doi: 10.1016/j.plantsci.2021.111063

Chen, Y., Wu, P., Zhao, Q., Tang, Y., Chen, Y., Li, M., et al. (2018). Overexpression of a phosphate starvation response AP2/ERF gene from physic nut in *Arabidopsis* alters root morphological traits and phosphate starvation-induced anthocyanin accumulation. *Front. Plant Sci.* 9:1186. doi: 10.3389/fpls.2018.01186

Cheng, M. C., Hsieh, E. J., Chen, J. H., Chen, H. Y., and Lin, T. P. (2012). Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with AtERF53 and negatively regulates the plant drought stress response. *Plant Physiol.* 158, 363–375. doi: 10.1104/pp.111.189738

Davies, P. J. (1987). *Plant Hormones and Their Role in Plant Growth and Development*. Netherlands: Springer. doi: 10.1007/978-94-009-3585-3

Dietz, K. J., Vogel, M. O., and Viehhauser, A. (2010). AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. *Protoplasma* 245, 3–14. doi: 10.1007/s00709-010-0142-8

Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M. (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC Box-mediated gene expression. *Plant Cell* 12, 393–404. doi: 10.1105/tpc.12.3.393

Gong, W., He, K., Covington, M., Dinesh-Kumar, S. P., Snyder, M., Harmer, S. L., et al. (2008). The development of protein microarrays and their applications in DNA–protein and protein–protein interaction analyses of *Arabidopsis* transcription factors. *Mol. Plant* 1, 27–41. doi: 10.1093/mp/ssm009

Gu, Y. Q., Wildermuth, M. C., Chakravarthy, S., Loh, Y. T., Yang, C., He, X., et al. (2002). Tomato transcription factors Pti4, Pti5, and Pti6 activate defense

responses when expressed in *Arabidopsis*. *Plant Cell* 14, 817–831. doi: 10.1105/tpc.000794

Guo, J., and Wang, M. H. (2011). Expression profiling of the DREB2 type gene from tomato (*Solanum lycopersicum* L.) under various abiotic stresses. *Hortic. Environ. Biotechnol.* 52, 105–111. doi: 10.1007/s13580-011-0125-5

Guo, Z. H., Zhang, Y. J., Yao, J. L., Xie, Z. H., Zhang, Y. Y., Zhang, S. L., et al. (2021). The NAM/ATAF1/2/CUC2 transcription factor PpNAC.A59 enhances PpERF.A16 expression to promote ethylene biosynthesis during peach fruit ripening. *Hortic. Res.* 8:209. doi: 10.1038/s41438-021-00644-6

Hirota, A., Kato, T., Fukaki, H., Aida, M., and Tasaka, M. (2007). The auxin-regulated AP2/EREBP gene PUCHI is required for morphogenesis in the early lateral root primordium of *Arabidopsis*. *Plant Cell* 19, 2156–2168. doi: 10.1105/tpc.107.050674

Huang, Z., Zhang, Z., Zhang, X., Zhang, H., Huang, D., and Huang, R. (2004). Tomato TERF1 modulates ethylene response and enhances osmotic stress tolerance by activating expression of downstream genes. *FEBS Lett.* 573, 110–116. doi: 10.1016/j.febslet.2004.07.064

Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M., Shima, Y., Nakamura, N., et al. (2017). Re-evaluation of the rin mutation and the role of RIN in the induction of tomato ripening. *Nat. Plants* 3, 866–874. doi: 10.1038/s41477-017-0041-5

Ito, Y., Nishizawa-Yokoi, A., Endo, M., Mikami, M., and Toki, S. (2015). CRISPR/Cas9-mediated mutagenesis of the RIN locus that regulates tomato fruit ripening. *Biochem. Biophys. Res. Commun.* 467, 76–82. doi: 10.1016/j.bbrc.2015.09.117

Jae-Hoon, L., Jong-Pil, H., Sang-Keun, O., Sanghyeob, L., Doil, C., and Woo Taek, K. (2004). The ethylene-responsive factor like protein 1 (CaERFLP1) of hot pepper (*Capsicum annuum* L.) interacts in vitro with both GCC and DRE/CRT sequences with different binding affinities: possible biological roles of CaERFLP1 in response to pathogen infection and high salinity conditions in transgenic tobacco plants. *Plant Mol. Biol.* 55, 61–81. doi: 10.1007/s11103-004-0417-6

Ji, A. J., Luo, H. M., Xu, Z. C., Zhang, X., Zhu, Y. J., Liao, B. S., et al. (2016). Genome-wide identification of the AP2/ERF gene family involved in active constituent biosynthesis in *Salvia miltiorrhiza*. *Plant Genome* 9, 1–11. doi: 10.3835/plantgenome2015.08.0077

Ji, S., Liu, Z., and Wang, Y. (2021). Trichoderma-induced ethylene responsive factor MsERF105 mediates defense responses in *Malus sieversii*. *Front. Plant Sci.* 12:708010. doi: 10.3389/fpls.2021.708010

Jia, M., Li, Y., Wang, Z., Tao, S., Sun, G., Kong, X., et al. (2021). TaIAA21 represses TaARF25-mediated expression of TaERFs required for grain size and weight development in wheat. *Plant J.* 108, 1754–1767. doi: 10.1111/tpj.15541

Jin, Y., Pan, W., Zheng, X., Cheng, X., Liu, M., Ma, H., et al. (2018). OsERF101, an ERF family transcription factor, regulates drought stress response in reproductive tissues. *Plant Mol. Biol.* 98, 51–65. doi: 10.1007/s11103-018-0762-5

Klay, I., Gouia, S., Liu, M., Mila, I., Khoudi, H., Bernadac, A., et al. (2018). Ethylene response factors (ERF) are differentially regulated by different abiotic stress types in tomato plants. *Plant Sci.* 274, 137–154. doi: 10.1016/j.plantsci.2018.05.023

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Lakehal, A., Dob, A., Rahneshan, Z., Novak, O., Escamez, S., Alallaq, S., et al. (2020). ETHYLENE RESPONSE FACTOR 115 integrates jasmonate and cytokinin signaling machineries to repress adventitious rooting in Arabidopsis. *New Phytol.* 228, 1611–1626. doi: 10.1111/nph.16794

Lang, Z., Wang, Y., Tang, K., Tang, D., Datsenka, T., Cheng, J., et al. (2017). Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit. *Proc. Natl. Acad. Sci. U.S.A.* 15, 4511–4519. doi: 10.1073/pnas.1705233114

Lee, J. H., Kim, D. M., Lee, J. H., Kim, J., Bang, J. W., Kim, W. T., et al. (2005). Functional characterization of NtCEF1, an AP2/EREBP-type transcriptional activator highly expressed in tobacco callus. *Planta* 222, 211–224. doi: 10.1007/s00425-005-1525-5

Lee, S. Y., Hwang, E. Y., Seok, H. Y., Tarte, V. N., Jeong, M. S., Jang, S. B., et al. (2015). Arabidopsis AtERF71/HRE2 functions as transcriptional activator via cis-acting GCC box or DRE/CRT element and is involved in root development through regulation of root cell expansion. *Plant Cell Rep.* 34, 223–231. doi: 10.1007/s00299-014-1701-9

Licausi, F., van Dongen, J. T., Giuntoli, B., Novi, G., Santaniello, A., Geigenberger, P., et al. (2010). HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana*. *Plant J.* 62, 302–315. doi: 10.1111/j.1365-313X.2010.04149.x

Lim, C., Kang, K., Shim, Y., Sakuraba, Y., An, G., and Paek, N. C. (2020). Rice ETHYLENE RESPONSE FACTOR 101 promotes leaf senescence through jasmonic acid-mediated regulation of OsNAP and OsMYC2. *Front. Plant Sci.* 11:1096. doi: 10.3389/fpls.2020.01096

Lin, R. C., Park, H. J., and Wang, H. Y. (2008). Role of *Arabidopsis* RAP2.4 in regulating light- and ethylene-mediated developmental processes and drought stress tolerance. *Mol. Plant* 1, 42–57. doi: 10.1093/mp/ssm004

Liu, M., Chen, Y., Chen, Y., Shin, J. H., Mila, I., Audran, C., et al. (2018). The tomato ethylene response factor Sl-ERF.B3 integrates ethylene and auxin signaling via direct regulation of Sl-Aux/IAA27. *New Phytol.* 219, 631–640. doi: 10.1111/nph.15165

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., et al. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell* 10, 1391–1406. doi: 10.1105/tpc.10.8.1391

Liu, Y., Zhao, T. J., Liu, J. M., Liu, W. Q., Liu, Q., Yan, Y. B., et al. (2006). The conserved Ala37 in the ERF/AP2 domain is essential for binding with the DRE element and the GCC box. *FEBS Lett.* 580, 1303–1308. doi: 10.1016/j.febslet.2006.01.048

Ma, Y., Zhang, F., Bade, R., Daxibater, A., Men, Z., and Hasi, A. (2015). Genome-wide identification and phylogenetic analysis of the ERF gene family in melon. *J. Plant Growth Regul.* 34, 66–77. doi: 10.1007/s00344-014-9443-z

McGrath, K. C., Dombrecht, B., Manners, J. M., Schenk, P. M., Edgar, C. I., Maclean, D. J., et al. (2005). Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. *Plant Physiol.* 139, 949–959. doi: 10.1104/pp.105.068544

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* 140:411. doi: 10.1104/pp.105.073783

Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-coffee: a novel method for multiple sequence alignments. *J. Mol. Biol.* 302, 205–217. doi: 10.1006/jmbi.2000.4042

Ouyang, Z. G., Liu, S. X., Huang, L. H., Hong, Y. B., Li, X. H., Huang, L., et al. (2016). Tomato SlERF.A1, SlERF.B4, SlERF.C3 and SlERF.A3, members of B3 group of ERF family, are required for resistance to *Botrytis cinerea*. *Front. Plant Sci.* 7:1964. doi: 10.3389/fpls.2016.01964

Park, J. M., Park, C. J., Lee, S. B., Ham, B. K., Shin, R., and Paek, K. H. (2001). Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2–type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. *Plant Cell* 13, 1035–1046. doi: 10.1105/tpc.13.5.1035

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Penin, A. A., Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., and Logacheva, M. D. (2019). Comparative analysis of developmental transcriptome maps of *Arabidopsis thaliana* and *Solanum lycopersicum*. *Genes* 10:50. doi: 10.3390/genes10010050

Pieterse, C. M. J., Leon-Reyes, A., Van der Ent, S., and Van Wees, S. C. M. (2009). Networking by small-molecule hormones in plant immunity. *Nat. Chem. Biol.* 5, 308–316. doi: 10.1038/nchembio.164

Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2002). DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem. Biophys. Res. Commun.* 290, 998–1009. doi: 10.1006/bbrc.2001.6299

Sakuma, Y., Maruyama, K., Qin, F., Osakabe, Y., and Yamaguchi-Shinozaki, K. (2006a). Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18822–18827. doi: 10.1073/pnas.0605639103

Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K., et al. (2006b). Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell* 18, 1292–1309. doi: 10.1105/tpc.105.035881

Solano, R., Stepanova, A., Chao, Q., and Ecker, J. R. (1998). Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes Dev.* 12, 3703–3714. doi: 10.1101/gad.12.23.3703

Song, C. P., Agarwal, M., Ohta, M., Guo, Y., Halfter, U., Wang, P. C., et al. (2005). Role of an *Arabidopsis* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. *Plant Cell* 17, 2384–2396. doi: 10.1105/tpc.105.033043

Song, X., Li, Y., and Hou, X. (2013). Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. pekinensis). *BMC Genomics* 14:573. doi: 10.1186/1471-2164-14-573

Sun, S., Yu, J. P., Chen, F., Zhao, T. J., Fang, X. H., Li, Y. Q., et al. (2008). TINY, a dehydration-responsive element (DRE)-binding protein-like transcription factor connecting the DRE- and ethylene-responsive element-mediated signaling pathways in *Arabidopsis*. *J. Biol. Chem.* 283, 6261–6271. doi: 10.1074/jbc.M706800200

Sun, Y. F., Liang, B., Wang, J., Kai, W. B., Chen, P., Jiang, L., et al. (2018). SlPti4 affects regulation of fruit ripening, seed germination and stress responses by modulating ABA signaling in tomato. *Plant Cell Physiol.* 59:1956. doi: 10.1093/pcp/pcy111

Tang, M., Sun, J., Liu, Y., Chen, F., and Shen, S. (2007). Isolation and functional characterization of the JcERF gene, a putative AP2/EREBP domain-containing transcription factor, in the woody oil plant *Jatropha curcas*. *Plant Mol. Biol.* 63:419. doi: 10.1007/s11103-006-9098-7

Tian, W., Huang, Y., Li, D., Meng, L., He, G., and He, T. (2021). Identification of StAP2/ERF genes of potato (*Solanum tuberosum*) and their multiple functions in detoxification and accumulation of cadmium in yest: implication for genetic-based phytoremediation. *Sci. Total Environ.* 810:152322. doi: 10.1016/j.scitotenv.2021.152322

Tournier, B., Sanchezballesta, M. T., Jones, B., Pesquet, E., Regad, F., Latché, A., et al. (2003). New members of the tomato ERF family show specific expression pattern and diverse DNA-binding capacity to the GCC box element. *FEBS Lett.* 550, 149–154. doi: 10.1016/S0014-5793(03)00757-9

Vainonen, J. P., Jaspers, P., Wrzaczek, M., Lamminmaki, A., Reddy, R. A., Vaahtera, L., et al. (2012). RCD1-DREB2A interaction in leaf senescence and stress responses in *Arabidopsis thaliana*. *Biochem. J.* 442, 573–581. doi: 10.1042/BJ20111739

Wang, L., Qin, L., Liu, W., Zhang, D., and Wang, Y. (2014). A novel ethylene-responsive factor from *Tamarix hispida*, ThERF1, is a GCC-box- and DRE-motif binding protein that negatively modulates abiotic stress tolerance in *Arabidopsis*. *Physiol. Plant* 152, 84–97. doi: 10.1111/ppl.12159

Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

Wei, G., Pan, Y., Lei, J., and Zhu, Y. X. (2005). Molecular cloning, phylogenetic analysis, expressional profiling and in vitro studies of TINY2 from *Arabidopsis thaliana*. *J. Biochem. Mol. Biol.* 38, 440–446. doi: 10.5483/BMBRep.2005.38.4.440

Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18, 691–699. doi: 10.1093/oxfordjournals.molbev.a003851

Wu, L., Zhang, Z., Zhang, H., Wang, X. C., and Huang, R. (2008). Transcriptional modulation of ethylene response factor protein JERF3 in the oxidative stress response enhances tolerance of tobacco seedlings to salt, drought, and freezing. *Plant Physiol.* 148, 1953–1963. doi: 10.1104/pp.108.126813

Xing, H. T., Jiang, Y. S., Zou, Y., Long, X. L., Wu, X. L., Ren, Y., et al. (2021). Genome-wide investigation of the AP2/ERF gene family in ginger: evolution and expression profiling during development and abiotic stresses. *BMC Plant Biol.* 21:561. doi: 10.1186/s12870-021-03329-3

Xu, S., Yao, S., Huang, R., Tan, Y., and Huang, D. (2020). Transcriptome-wide analysis of the AP2/ERF transcription factor gene family involved in the regulation of gypenoside biosynthesis in *Gynostemma pentaphyllum*. *Plant Physiol. Biochem.* 154, 238–247. doi: 10.1016/j.plaphy.2020.05.040

Xu, Z. S., Xia, L. Q., Chen, M., Cheng, X. G., Zhang, R. Y., Li, L. C., et al. (2007). Isolation and molecular characterization of the *Triticum aestivum* L. ethylene-responsive factor 1 (TaERF1) that increases multiple stress tolerance. *Plant Mol. Biol.* 65, 719–732. doi: 10.1007/s11103-007-9237-9

Yamasaki, K., Kigawa, T., Seki, M., Shinozaki, K., and Yokoyama, S. (2013). DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. *Trends Plant Sci.* 18, 267–276. doi: 10.1016/j.tplants.2012.09.001

Yang, J., Liu, Y., Yan, H., Tian, T., You, Q., Zhang, L., et al. (2018). PlantEAR: functional analysis platform for plant EAR motif-containing proteins. *Front. Genet.* 9:590. doi: 10.3389/fgene.2018.00590

Yang, Y., Zhu, G., Li, R., Yan, S., Fu, D., Zhu, B., et al. (2017). The RNA editing factor SlORRM4 is required for normal fruit ripening in tomato. *Plant Physiol.* 175, 1690–1702. doi: 10.1104/pp.17.01265

Zhang, G., Chen, M., Li, L., Xu, Z., Chen, X., Guo, J., et al. (2009). Overexpression of the soybean GmERF3 gene, an AP2/ERF type transcription factor for increased tolerances to salt, drought, and diseases in transgenic tobacco. *J. Exp. Bot.* 60, 3781–3796. doi: 10.1093/jxb/erp214

Zhang, H., Huang, Z., Xie, B., Chen, Q., Tian, X., Zhang, X., et al. (2004). The ethylene-, jasmonate-, abscisic acid- and NaCl-responsive tomato transcription factor JERF1 modulates expression of GCC box-containing genes and salt tolerance in tobacco. *Planta* 220, 262–270. doi: 10.1007/s00425-004-1347-x

Zhang, Y., Ji, A., Xu, Z., Luo, H., and Song, J. (2019). The AP2/ERF transcription factor SmERF128 positively regulates diterpenoid biosynthesis in *Salvia miltiorrhiza*. *Plant Mol. Biol.* 100, 83–93. doi: 10.1007/s11103-019-00845-7

Zhang, Z., Zhang, H., Quan, R., Wang, X. C., and Huang, R. (2009). Transcriptional regulation of the ethylene response factor LeERF2 in the expression of ethylene biosynthesis genes controls ethylene production in tomato and tobacco. *Plant Physiol.* 150, 365–377. doi: 10.1104/pp.109.135830

Zhu, Y., Li, Y., Zhang, S., Zhang, X., Yao, J., Luo, Q., et al. (2018). Genome-wide identification and expression analysis reveal the potential function of ethylene responsive factor gene family in response to Botrytis cinerea infection and ovule development in grapes (*Vitis vinifera* L.). *Plant Biol.* 21, 571–584. doi: 10.1111/plb.12943

Zong, Y. X., Hao, Z. Y., Tu, Z. H., Shen, Y. F., Zhang, C. G., Wen, S. Y., et al. (2021). Genome-wide survey and identification of AP2/ERF genes involved in shoot and leaf development in *Liriodendron* chinense. *BMC Genomics* 22:807. doi: 10.1186/s12864-021-08119-7

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 07 March 2022
doi: 10.3389/fpls.2022.833747

ARF6s Identification and Function Analysis Provide Insights Into Flower Development of *Punica granatum* L.

*Yujie Zhao[1,2], Yuying Wang[1,2], Xueqing Zhao[1,2], Ming Yan[1,2], Yuan Ren[1,2] and Zhaohe Yuan[1,2]**

[1] Co-innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China, [2] College of Forestry, Nanjing Forestry University, Nanjing, China

Based on the genome and small-RNA sequencing of pomegranate, miRNA167 and three target genes *PgARF6* were identified in "Taishanhong" genome. Three *PgARF6* genes and their corresponding protein sequences, expression patterns in pomegranate flower development and under exogenous hormones treatments were systematically analyzed in this paper. We found that PgARF6s are nuclear proteins with conserved structures. However, PgARF6s had different protein structures and expression profiles in pomegranate flower development. At the critical stages of pomegranate ovule sterility (8.1–14.0 mm), the expression levels of *PgARF6s* in bisexual flowers were lower than those in functional male flowers. Interestingly, *PgARF6c* expression level was significantly higher than *PgARF6a* and *PgARF6b*. Under the treatment of exogenous IBA and 6-BA, *PgARF6s* were down-regulated, and the expression of *PgARF6c* was significantly inhibited. PgmiR167a and PgmiR167d had the binding site on *PgARF6* genes sequences, and PgARF6a has the directly targeted regulatory relationship with PgmiR167a in pomegranate. At the critical stage of ovule development (8.1–12.0 mm), exogenous IBA and 6-BA promoted the content of GA and ZR accumulation, inhibited BR accumulation. There was a strong correlation between the expression of *PgARF6a* and *PgARF6b*. Under exogenous hormone treatment, the content of ZR, BR, GA, and ABA were negatively correlated with the expressions of *PgARF6* genes. However, JA was positively correlated with *PgARF6a* and *PgARF6c* under IBA treatment. Thus, our results provide new evidence for *PgARF6* genes involving in ovule sterility in pomegranate flowers.

Keywords: pomegranate, auxin response factor 6, expression analysis, PgmiR167, endogenous hormones, flower development

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Qiang Xu,
Huazhong Agricultural University, China
Lailiang Cheng,
Cornell University, United States

***Correspondence:**
Zhaohe Yuan
zhyuan88@hotmail.com

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

Received: *12 December 2021*
Accepted: *26 January 2022*
Published: *07 March 2022*

Citation:
Zhao Y, Wang Y, Zhao X, Yan M, Ren Y and Yuan Z (2022) ARF6s Identification and Function Analysis Provide Insights Into Flower Development of Punica granatum L. Front. Plant Sci. 13:833747. doi: 10.3389/fpls.2022.833747

INTRODUCTION

Pomegranate, belonging to Lythraceae (Yuan et al., 2018), is rich in polyphenols such as punicalagin, ellagic acid, and other ellagitannin-based compounds (Johanningsmeier and Harris, 2011). Pomegranate is an excellent fruit tree with economic, nutritional, medicinal, ornamental, and ecological values (Yuan, 2015), therefore it is reputed as a "super fruit" (Teixeira da Silva et al., 2013). Pomegranate trees produce bisexual flower and functional male flower. Bisexual flower has an apparent tubular shape, and can bear a fruit as time advancement, whereas functional male

flower is bell-shaped which usually drops and fails to set fruit. Studies have revealed that female sterility in functional male flowers is due to the abnormal ovule development failing to generate egg cells (Cai et al., 1993; Chen L. N. et al., 2017). When the vertical diameter of bisexual flowers is 5.1–10.0 mm, inner integument primordia are being form. When it is 10.1–13.0 mm, the outer integument primordia are formed in bisexual flowers, while the inner integument primordia are characterized in functional male flowers. When bisexual flowers gain their vertical diameter to 13.1–15.0 mm, the integument continues to enlarge, whereas the inner integument of functional male flowers is stagnated and the ovules display wilting (Chen L. N. et al., 2017). It can be concluded that it is an important stage for pomegranate ovule development when the vertical diameter of a flower bud is 5.1–15.0 mm.

Auxin is closely related to the differentiation of plant cells and vascular tissues, the morphogenesis of rhizomes, apical dominance, tropism, and response to external pressure (Guilfoyle and Hagen, 2007; Kumar et al., 2011; Yang et al., 2013). Auxin is a significant growth regulator which participates in the growth and development process of plants by changing the expression of early reactive genes, such as *Aux*/*IAA*, *GH3,* and *SAUR* gene families (Abel and Theologis, 1996; Wan et al., 2014; Roosjen et al., 2017). The promoter region of these genes contain one *cis*-acting element (TGTCTC) consisting of six bases, which is known as auxin response elements (AuxRE) (Guilfoyle et al., 1998). ARF (Auxin response factor) is a transcription factor that can specifically recognize and bind auxin response elements. ARF participates in auxin signal transduction by activating or inhibiting auxin response gene expression (Tiwari et al., 2003; Zhao et al., 2016; Zhang F. et al., 2020). Previous studies have shown that ARFs play independent roles in auxin signal transduction, and ARFs interact with *WUS*, *AG,* and *AP2* which are involved in regulating the morphogenesis of pistillate primordia (Liu X. G. et al., 2014). In *Arabidopsis thaliana* ARF family members, *AtARF1* and *AtARF2* are involved in regulation of pistil growth, flower organ development and wither (Ulmasov et al., 1997; Ellis et al., 2005). *AtARF3* and *AtARF4* participate in plant reproduction and vegetative growth (Sessions et al., 1997; Pekker et al., 2005). *AtARF5* plays a critical role in the formation of vascular tissue and the establishment of embryo development model (Hardtke and Berleth, 1998; Johnson and Douglas, 2007; Miyashima et al., 2013). *AtARF6* and *AtARF8* are suggested to be involved in regulating floral organ development and affect ovule development (Tian et al., 2004; Nagpal et al., 2005). *ARF6* and *ARF8* promote jasmonic acid (JA) production and flower maturation (Nagpal et al., 2005). *AtARF8* is involved in regulating hypocotyl elongation and root growth (Tian et al., 2004; Nagpal et al., 2005; Wu et al., 2006). *ATARF7*/*ATARF19*, *ATIAA1,* and *ATLBD16*/*18*/*29* are involved in the regulation network of lateral root development in *Arabidopsis thaliana* (Fukaki et al., 2005). In the *arf7*/*arf19*, the overexpression of *LBD16/ASL18* and *LBD29/ASL16* lead to the formation of lateral roots (Fukaki et al., 2005; Li et al., 2006; Okushima et al., 2007; Feng et al., 2012). In "Dabenzi" pomegranate genome, three *ARF6* homologous genes are identified, which might result from *ARF6* functional differentiation (Huang et al., 2019).

MiRNAs (microRNAs) are the class of important non-coding small RNAs, the length of which is 21–24 nt and play essential roles in regulating plant growth and development. The previous results demonstrate that miRNAs regulate the expressions of *ARF* family members. The expressions of *ARF10*, *ARF16,* and *ARF17* are regulated by miR160 (Damodharan et al., 2016), and the target genes of miR167 are *ARF6* and *ARF8* (Wu et al., 2006). miR167, together with *ARF6* and *ARF8*, is involved in regulating the development and maturity of pistil and stamen (Wu et al., 2006). Previous studies have shown that miR167 overexpression leads to reduced ovule maturity (Ru et al., 2006). By degrading *ARF8* mRNA and inhibiting translation of *ARF6* mRNA, miR167 regulate their expression and participate in the plant's growth and development. In fact, miR167 define the expression regions of *ARF6* and *ARF8* in stamens and ovules, which ensure the normal development of sepals, petals, anthers, pistils, and ovules (Ru et al., 2006; Wu et al., 2006). Up to now, the roles of *ARF6* have been investigated in *Arabidopsis thaliana*, tomato (Tang, 2016), apple (Li H. F. et al., 2018), and strawberry (Wang, 2019). However, there is no relevant report on *ARF6* function in pomegranate flowers development. By integrating the genomic, transcriptomic and small-RNA sequencing data of "Taishanhong" pomegranate, *PgARF6s* were identified, and the targeted relationships between *ARF6* homologous genes and miR167 were preliminarily demonstrated in pomegranate. At the same time, the correlation between the expression of *PgARF6s* and endogenous hormone content was analyzed during pomegranate flower development. Our results laid the foundation for clarifying the function of *ARF6* in pomegranate flower development.

MATERIALS AND METHODS

Plant Material and Treatments

Functional male flowers and bisexual flowers of pomegranate were categorized into eight stages according to their pistil vertical diameter (Zhao et al., 2020), that is 3.0–5.0 mm (P1), 5.1–8.0 mm (P2), 8.1–10.0 mm (P3), 10.1–12.0 mm (P4), 12.1–14.0 mm (P5), 14.1–16.0 mm (P6), 16.1–18.0 mm (P7), and 18.1–20.0 mm (P8). The flower buds were collected at the Baima Base for Teaching and Scientific Research of Nanjing Forestry University. Three biological repeats of samples from each stage were prepared, frozen with liquid nitrogen and stored in the −78°C refrigerator for subsequent use.

Nine six-year-old "'Taishanhong"' plants were treated with 20.0 mg/L IBA and 100.0 mg/L 6-BA, respectively (Chen et al., 2020). After 1one week, functional male flowers and bisexual flowers were collected as test samples. The control samples were treated with clear water. The determination method of endogenous hormone content was described elsewhere (Zhao et al., 2021).

Identification and Bioinformatics Analysis of PgARF6s in Pomegranate

Firstly, using the published "Taishanhong" pomegranate genome data (Yuan et al., 2018) and ARF (PF06507) data in the Pfam

database, the hmmbuild model was adopted in the HMMER 3.0 software package, and the hmmsearch program was further used to screen candidate ARF protein sequences (*e*-value $< 1e^{-5}$, identity $> 50\%$). ARF protein sequences in plant transcription factor database were used as queries to perform BLAST against the pomegranate genome database (*e*-value $< 1e^{-5}$, identity $> 50\%$), duplication was removed and ARF protein sequences were screened and selected. Then, the CDD[1] and SMART[2] were used to detect protein structure domains of the candidate sequences. The exon-intron structures of *PgARF6* genes were displayed using Gene Structure Display Server 2.0.[3] Motifs of *PgARF6s* nucleic acids sequences were identified using MEME[4] with default parameter. Phyre 2.0 was used to predict ARF6 proteins structures.

Eucalyptus grandis (Eucgr), *Populus euphratica* (Pop), and *Carica papaya* (Capa) protein sequences were downloaded from plant transcription factor database, and *Arabidopsis thaliana* (AT) protein sequences were downloaded from TAIR database. The ARF phylogenetic tree of candidate proteins of *Eucalyptus grandis*, *Populus euphratica*, *Carica papaya*, and *Arabidopsis thaliana* were constructed on the online website.[5] The amino acid sequences of PgARF6s were aligned by ClustalX. The evolutionary tree of species was constructed with MEGA 7.0 software, and plants protein sequences were downloaded from NCBI database.

PgARF6 proteins subcellular localization was predicted by Cell-PLoc 2.0.[6] At the same time, the String (version 11.5) was used to predict the PgARF6 plant proteins interaction network, and the number of organism was set as "*Arabidopsis thaliana*."

To examine expression patterns of *PgARF6s* in different pomegranate tissues, the published transcriptome data (Chen L. N. et al., 2017; Qin et al., 2017) were download for expression analysis (**Supplementary Table 1**). Firstly, all RNA-Seq data were qualitatively controlled by fastp to obtain cleaned reads. Then, the pomegranate transcriptome data were analyzed according to the method reported by Zhao et al. (2020). Finally, the heatmap of *PgARF6s* transcription levels was drawn.

RNA Extraction, cDNA Synthesis and Gene Cloning

Total RNAs were extracted from the flower buds of pomegranate using the BioTeke plant total RNA extraction kit. The integrity and concentration of RNA were detected by agar-gel electrophoresis and UV-visible spectrophotometer ND-2000 (Thermo Fisher Scientific Inc., United States). cDNA was prepared by reverse transcription kit (PrimeScript^{TM}RT reagent Kit with gDNA Eraser, TaKaRa).

The CDS-full-length of *PgARF6a*, *PgARF6b,* and *PgARF6c* were cloned by homologous cloning. The CDS of *PgARF6* genes were extracted from the genome data of "Taishanhong" pomegranate, and Oligo 7.0 was used to design the gene clone primers (**Table 1**). A high-fidelity PCR reaction was carried out using the cDNA of mixed flowers samples as a template. The reaction conditions were as follows: first 98°C for 2 min; then 98°C for 15 s, 58°C for 15 s, 72°C for 1 min, and 30 cycles; last 72°C for 5 min. Then, the recovered PCR products were sent to Shanghai Sangon Biological Company for sequencing.

PgARF6s Subcellular Localization

The linearized pBI121 vector was obtained by double digestion of restriction enzymes *Xba*I and *Bam*HI. Homologous arm primers were designed by adding restriction site sequences at both ends of the target genes (**Table 1**). According to ClonExpress®II One Step Cloning Kit (Vazyme Biotech Co., Ltd.), *PgARF6s* were connected with the modified pBI121-GFP vector, and the recombinant vectors were transferred into *DH5*α. Then, the constructed vectors plasmids were extracted and transferred into Agrobacterium GV1301 competent cells by the freeze-thaw method.

Monoclonal colonies were selected and inoculated into a 20 mL liquid medium. The bacteria solution was 5,000 r/min, centrifuged for 5 min to collect the bacteria, and then added with 20 mL leaching solution to re-suspend the bacteria. The suspension bacteria were injected into tobacco. After injection, tobacco returned to normal growth after dark culture for 24 h, and fluorescence observation was performed 1–2 days later.

Expression Analysis of PgARF6s

The expression patterns of *PgARF6s* were studied using bisexual flowers and functional male flowers from different development stages. According to the CDS of *PgARF6s*, quantitative realtime PCR (qRT-PCR) primers were designed (**Table 1**). qRT-PCR methods had been described previously (Zhao et al., 2020), and the data were analyzed by $2^{-\Delta\Delta CT}$ method.

Prediction of PgmiR167 Target Sites of PgARF6s in Pomegranate

The nucleotide sequences of *PgARF6s* were analyzed by psRNATarget online software to predict the target site of PgmiR167.

Functional male flowers and bisexual flowers with the vertical diameter of 5.0–10.0, 10.1–13.0, and 13.1–18.0 mm were selected and sent to Novogene Company (Beijing) for transcriptome (PRJNA754480) and small-RNA sequencing (PRJNA793612). Sequencing libraries were generated using NEBNext® Ultra^{TM} RNA Library Prep Kit for Illumina® (NEB, United States). Small RNA libraries were generated using Small RNA Sample Pre Kit. Index of pomegranate reference genome (ASM220158v1) was built using Hisat2 v2.0.5 and clean reads were aligned to pomegranate genome using Hisat2 v2.0.5. featureCounts v1.5.0-p3 was used to count the reads numbers mapped to each gene. Then, FPKM of each gene was calculated based on the length of the gene and reads count mapped to this gene. FPKMs of *PgARF6s* and PgmiR167 were converted into $\mathrm{Log}_2(\mathrm{FPKM}+1)$.

[1] https://www.ncbi.nlm.nih.gov/cdd

[2] http://smart.embl-heidelberg.de//

[3] http://gsds.cbi.pku.edu.cn

[4] https://meme-suite.org/meme/tools/meme

[5] https://www.omicstudio.cn/tool/53

[6] http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/

TABLE 1 | Primers for the gene cloning, subcellular localization, and expression analysis.

Primer	Primer sequence	Annotation	Length of product
PgARF6a	F: ATGAGGCTTTCTTCTGCTT; R: TCAGTAATCGAGGGACCCC	Gene clone	2,700
PgARF6b	F: ATGAGACTGTCTCCAGCTG; R: TTAGTAATCAAGAGATCCT	Gene clone	2,901
PgARF6c	F: ATGAGGCTCTCTTCGCCTG; R: TCAGTAGTCAAGAGACCCT	Gene clone	2,607
PgARF6a	F: GAGAACACGGGGGACTCTAGAATGAGGCTTTCTTCTGCTT; R: GCCCTTGCTCACCATGGATCCGTAATCGAGGGACCCCACG	Subcellular localization	2,700
PgARF6b	F: GAGAACACGGGGGACTCTAGAATGAGACTGTCTCCAGCTG; R: GCCCTTGCTCACCATGGATCCGTAATCAAGAGATCCTCTT	Subcellular localization	2,901
PgARF6c	F: GAGAACACGGGGGACTCTAGAATGAGGCTCTCTTCGCCTG; R: GCCCTTGCTCACCATGGATCCGTAGTCAAGAGACCCTATG	Subcellular localization	2,607
qRTPgARF6a	F: CCTCCTTTGTTCTTTGCC; R: CATCTTGGGTGAGACGTTAT	Gene expression	200
qRTPgARF6b	F: CTGTCGGAACCCGTGTAGT; R: CAGGGTCATCTGAGCATAAA	Gene expression	200
qRTPgARF6c	F: CAACCCAAGGGCAAGTCC; R: AGCATCCTGAACCGCATT	Gene expression	200
PgActin	F: AGTCCTCTTCCAGCCATCTC; R: CACTGAGCACAATGTTTCCA	Gene expression	200

Statistical Analysis

All experimental data were analyzed using SPSS software (22.0, United States). The data were expressed as the means and standard deviations (mean ± SD). Different letters were the significant difference among developmental stages for flower, based on Duncan's Test at the $p < 0.05$ level.

RESULTS

Identification and Bioinformatics Analysis of Pomegranate ARF6

A total of 19 ARF genes were identified from the genome of "Taishanhong" pomegranate using two methods of local blast and hmmsearch model. All candidate proteins were confirmed to belong to ARF family (**Supplementary Table 2**). CDD and SMART conserved domains showed that Pg006237.1, Pg013200.1, and Pg026615.1 have typical B3, Auxin_resp, and Aux_IAA conserved domains. Pg006237.1, Pg013200.1, and Pg026615.1 were renamed as PgARF6a, PgARF6b, and PgARF6c for convenience in subsequent analysis, based on their potential orthologs proteins in *Arabidopsis thaliana* (**Figure 1**).

Gene structures were investigated on gene structure display server (GSDS). *PgARF6a*, *PgARF6b,* and *PgARF6c* contained 14 exons, indicated that the gene structure with 13 introns was the main form of *PgARF6s*. The results of conservative motifs showed similar conservative motif in *PgARF6* genes. These results indicated that the motif and gene structure of *PgARF6s* were highly consistent and conservative (**Figure 2**).

To better understand protein structure and functional information of pomegranate ARF6s, we performed a systematic bioinformatics analysis of PgARF6 protein sequences. The multi-sequence alignment analysis showed that three PgARF6 homologous proteins had long conserved sequences in amino acid composition, but protein sequences were different at the sites 387–607 (PgARF6a) (**Figure 3A**). Proteins insertions were observed at sites 387–389 (AMK) of PgARF6c and sites 492–519 (FVQGFPESQASAQAQAQAQLLQQQLQRQ) of PgARF6b, and proteins deletions were found at sites 419–421 (LGL) of PgARF6c and sites 678–683 (SSPSSW) of PgARF6b. Then, tertiary structure analysis of PgARF6 proteins showed that the structure of ARF6 proteins were highly similar, but the 3_10_Helix (red) structures were located differently in PgARF6a, PgARF6b, and PgARF6c (**Figure 3B**). We found that the tertiary structure of PgARF6b and PgARF6c were derived from PgARF6a, and the structure was also different between PgARF6b and PgARF6c.

The Cell-PLoc 2.0 prediction results showed that PgARF6a, PgARF6b, and PgARF6c were located in the cell nucleus. CDSs of *PgARF6a*, *PgARF6b*, and *PgARF6c* were cloned (**Supplementary Figure 1**) and submitted to the NCBI database (GenBank number: OM179998, OM179999, OM180000). Then, GFP fluorescence signals of the control were observed in both cell membrane and nucleus, while that of PgARF6-GFP were only observed in the nucleus (**Figure 4**), indicating that PgARF6 proteins were localized in cell nucleus and had the subcellular localization characteristics of typical transcription factors.

Protein-protein interaction analysis results (**Supplementary Figure 2**) showed that PgARF6a, PgARF6b, and PgARF6c had the highest similarity with AtARF6 (*E*-value $0.0e^{+00}$, $2.2e^{-305}$, $6.1e^{-302}$, respectively), suggesting that PgARF6s might have the same function with AtARF6. ARF6 with Aux/IAA proteins fuse heterodimers that is involved in regulating auxin response gene expression, stamen and gynoecium maturation (Piya et al., 2014). ARF6 promotes JA production, partially redundant with ARF8. ARF6, BZR1, and BEE 2 positively regulate brassinosteroid (BR) signaling (Li Q. F. et al., 2018; Tian et al., 2018). IAA8, IAA19, and IAA28 are short-lived transcriptional factors that inhibit the expression of auxin response genes at low auxin concentrations (Piya et al., 2014). The repressors are due to their binding to AuxRE in the ARF promoter region (Hussain et al., 2020).

To understand the relationships between ARF6 proteins in pomegranate and other plants, we conducted phylogenetic analyses of known *ARF6* using amino acid sequences. It was found that PgARF6 proteins were highly similar with other plants, indicating that the amino acid sequences of ARF6 in different plants were relatively conservative (**Supplementary Figure 3**). Phylogenetic analyses showed

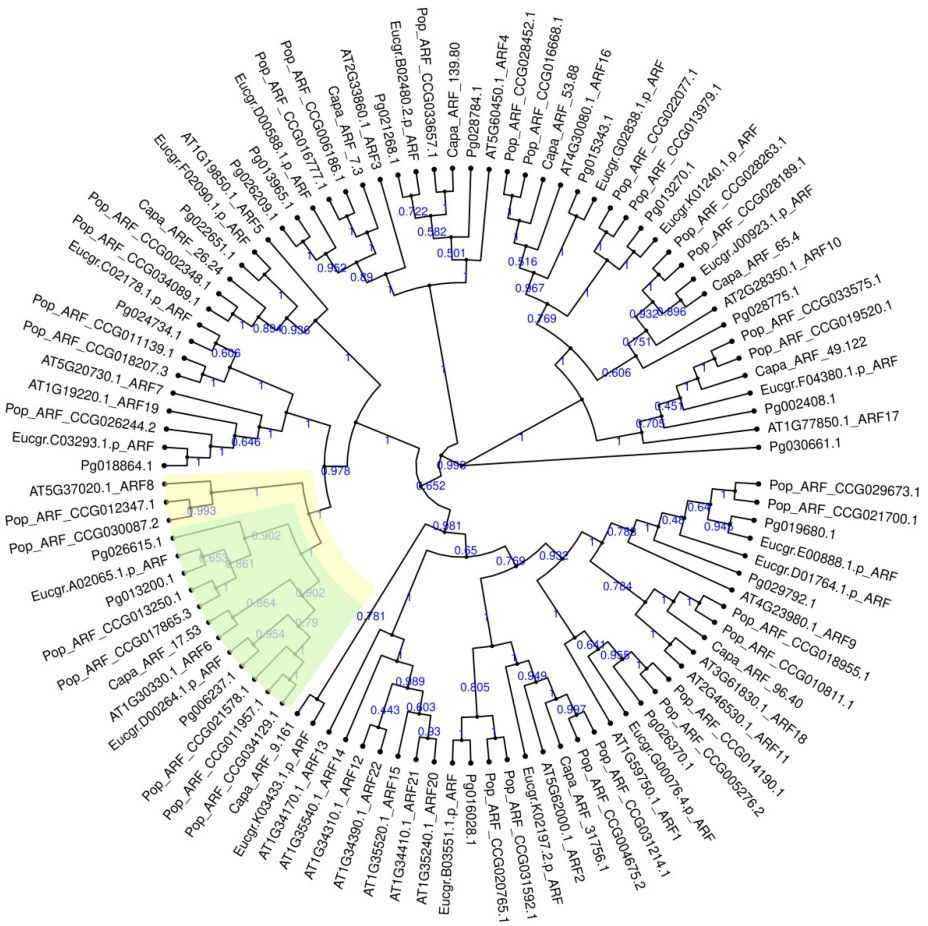


FIGURE 1 | Phylogenetic tree of ARF gene family in *Punica granatum* (Pg), *Eucalyptus grandis* (Eucgr), *Populus euphratica* (Pop), *Carica papaya* (Capa), and *Arabidopsis thaliana* (AT).

that PgARF6s could successfully cluster with *Arabidopsis thaliana*, *Solanum lycopersicum*, *Citrus clementina*, and *Carica papaya* (**Supplementary Figure 4**). PgARF6 proteins had a far evolutionary relationship with ARF6 of *Zea mays*, *Ananas comosus,* and *Malus domestica*, whereas PgARF6 proteins clustered more closely to *Solanum lycopersicum*, *Morus notabilis,* and *Carica papaya*.

In order to uncover the roles of *PgARF6* genes in pomegranate different organs, *PgARF6s* expression levels in different tissues were analyzed using transcriptome data. The results showed that *PgARF6s* were expressed in leaf, root, flower, fruit, endocarp, and pericarp (**Figure 5**). However, the results revealed the differences of expression levels for three *PgARF6* genes. We found that *PgARF6s* were weakly expressed in seed coat, whereas *PgARF6s* were highly expressed in flowers, indicating that *PgARF6s* might be involved in pomegranate flower organ development.

Expression Profiles of PgARF6s During Flower Development in Pomegranate

To explore the putative function of *PgARF6s* in floral organ development, the expression patterns of *PgARF6s* in bisexual flowers and functional male flowers were analyzed by qRT-PCR (**Figure 6**). The expression level of *PgARF6a* in bisexual flowers was the highest at P6, which was 9 times than that at P5 and 4.6 times than that at P1. At the critical stage of ovule primordia formed (5.1–8.0 mm, P2), *PgARF6a* transcriptional level in bisexual flowers was significantly higher than that in functional male flowers. During flowers development, the expression pattern of *PgARF6b* was consistent with that of *PgARF6a*. At the critical stage of ovule sterility (8.1–12.0 mm, P3–P4), transcriptional levels of *PgARF6a* and *PgARF6b* in bisexual flowers were significantly lower than that in functional male flowers.

The expression level of *PgARF6c* in bisexual flowers was the highest at the P2 stage, which were 9.5 times than that of P3 and 15 times than that of P7. During functional male flower development stages, *PgARF6c* transcriptional level was the highest at the P4 stage, which was 7.1 times higher than that at P5 stage. At the period of ovule primordia formed (5.1–8.0 mm, P2), the expression levels of *PgARF6a*, *PgARF6b,* and *PgARF6c* in bisexual flowers were higher compared with that in functional male flowers. These results showed that *PgARF6s* expression levels in bisexual flowers were significantly lower than that in

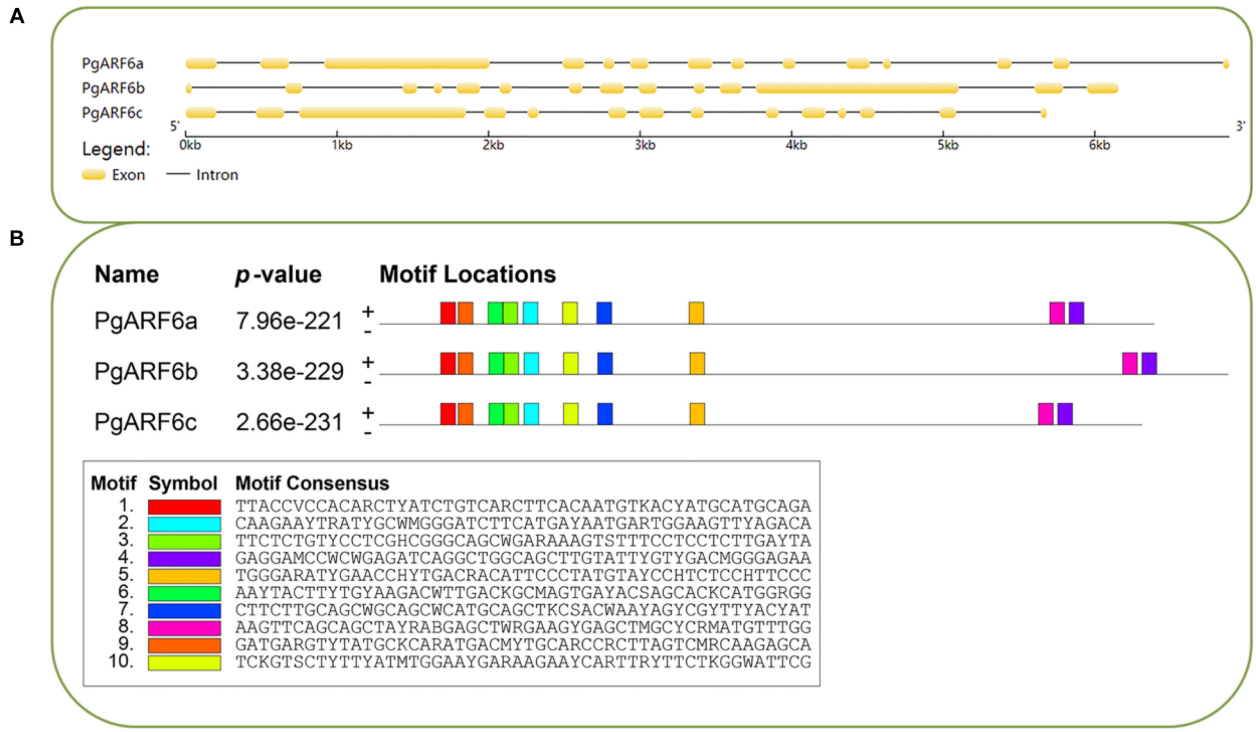


FIGURE 2 | Basic information of *PgARF6* sequences in pomegranate. **(A)** Gene structure of *PgARF6s*, **(B)** conservative nucleic acid motifs of *PgARF6s*.

functional male flowers at the critical stage of ovule sterility (8.1–12.0 mm, P3–P4), indicating that *PgARF6s* was involved in ovule sterility developing into functional male flower.

The Relationships of PgmiR167 Targeted PgARF6s in Pomegranate

To investigate the targeted relationships between *PgARF6s* and PgmiR167, we predicted the target sites of PgmiR167 based on the *PgARF6* sequences (**Table 2**). There were two base differences between PgmiR167a and PgmiR167d that A of PgmiR167a was replaced by GG. PgmiR167a had the same sequence of binding sites on *PgARF6*, and there was only one binding site. PgmiR167d and PgmiR167a had a similar situation. The results indicated that *PgARF6s* genes were targets of PgmiR167a and PgmiR167d, suggesting that regulation of ARF6s by miR167 was conserved between pomegranate, tomato (Liu N. et al., 2014) and *Arabidopsis* (Wu et al., 2006).

To determine the relationships between PgmiR167 and *PgARF6* genes at transcript level, we further analyzed the expression profiles of *PgARF6* and PgmiR167 based on data of transcriptome and small-RNA sequencing (**Figure 7**). During functional male flower development, the expression level of *PgARF6a* increased, but PgmiR167a expression level gradually decreased. During bisexual flowers development, the expression of *PgARF6a* gradually increased, whereas the expression of PgmiR167a decreased. The transcriptional level of PgmiR167a gradually decreased in functional male flowers and bisexual flowers, while the expression level of PgmiR167d decreased then increased. The expression levels of *PgARF6b* slowly decreased in bisexual flowers and functional male flowers, while an opposite trend of *PgARF6a* was observed. The transcriptional level of *PgARF6c* was firstly increased and then decreased in functional male flowers, while gradually increased in bisexual flowers. These results indicated that PgmiR167a directly regulated *PgARF6a* in pomegranate flowers.

Expression of PgARF6 Genes in Response to 6-BA Treatment

To determine the response of PgARF6 genes to exogenous cytokinin stimuli, their expression patterns in pomegranate flowers were investigated after 6-BA treatments (**Figure 8**). Under 6-BA treatment, the expression level of *PgARF6a* in bisexual flowers was higher than that in functional male flowers at the critical stage of ovule development (3.0–12.0 mm, P1–P4). The expression level of *PgARF6a* in bisexual flowers was 1.9 times and 1.7 times higher than that in functional male flowers at P4 and P3 stage, respectively. Whereas *PgARF6a* expression level in bisexual flowers was lower than that in functional male flowers at the development period of organ maturation (P5–P8). After being induced by exogenous 6-BA, *PgARF6b* transcriptional level in bisexual flowers was lower than that in functional male flowers at the period of ovule primordia development (3.0–8.0 mm, P1–P2). At the critical stage of ovule sterility (8.1–12.0 mm, P3–P4), *PgARF6b* transcriptional level in bisexual flowers was higher than that in functional male flowers. The transcriptional level of *PgARF6b* in bisexual flowers was 3.7, 2.1, and 2.5 times

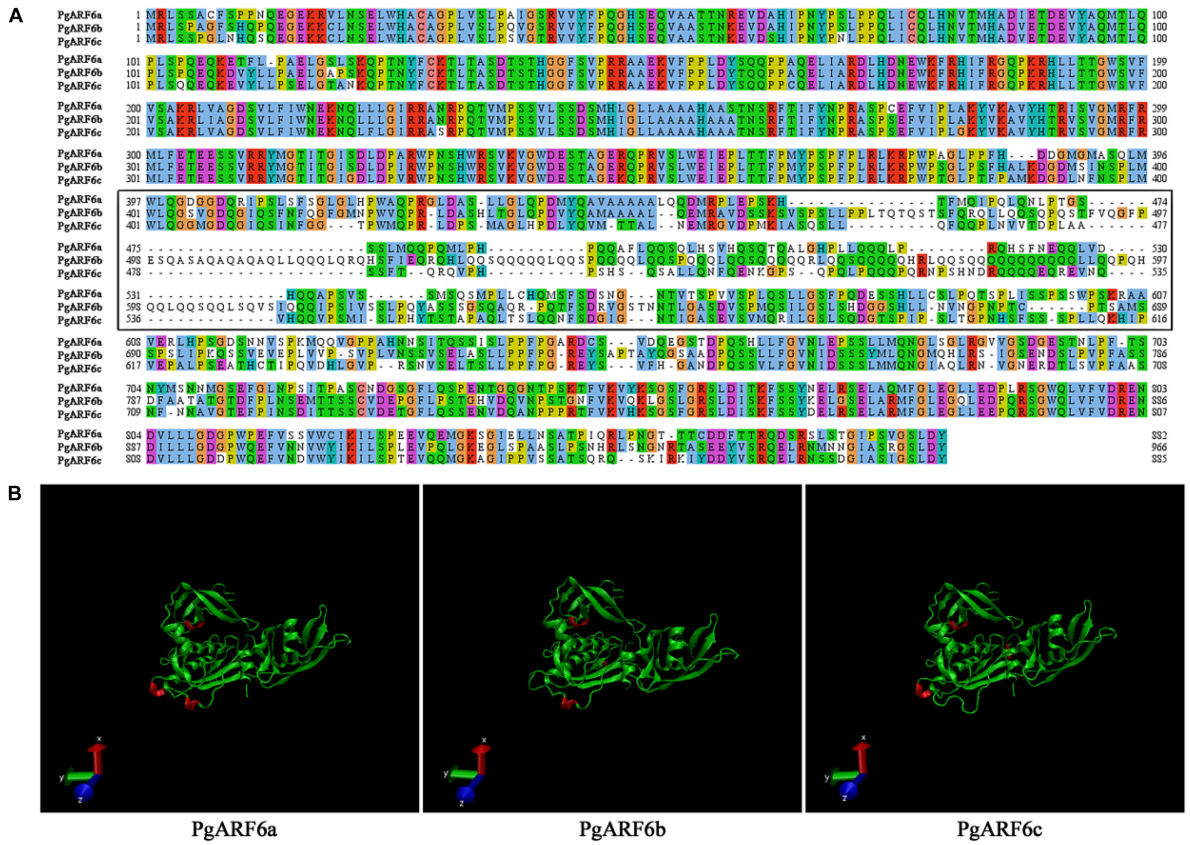


FIGURE 3 | (A) Sequences alignment of PgARF6s. **(B)** The tertiary structure of PgARF6s.

higher than that in functional male flowers at P3, P4, and P5, respectively.

Under 6-BA treatment, *PgARF6c* expression level in bisexual flowers was higher than that in functional male flowers at the critical stage of ovule development (5.1–12.0 mm, P2–P4). *PgARF6c* expression level in bisexual flowers was 2.9 times higher than that in functional male flowers at P2. The level of *PgARF6c* was down-regulated after P2 period.

Response of Endogenous Hormone in Pomegranate Flowers to Exogenous 6-BA Treatment

The content of IAA under exogenous 6-BA treatment decreased, increased and then decreased during bisexual flowers development (**Figure 9**). The content of IAA in the treatment was higher than that in the control at P2 and P4–P5 stages of the critical stage of ovules development. During functional male flowers development, the content of IAA induced by exogenous 6-BA was higher than that in the control, but lower than that in the control at P1.

Under 6-BA treatment, ZR content in bisexual flowers gradually increased, but there were no significant differences among P6, P7, and P8. At the critical stage of ovule development in bisexual flower (P1–P3), the content of IAA in the treatment was higher compared with that in the control. During the development of functional male flowers, the change pattern of ZR content in the treatment was consistent with the control. During P1–P3 stages of functional male flower, the content of IAA in the treatment was lower than that in the control. Under 6-BA treatment, the content of ZR in bisexual flowers was lower than that in functional male flowers at P1–P4 stages, which was consistent with the control. However, the content of ZR in bisexual flowers was higher than that in functional male flowers at P5–P8 stages (**Figure 9**).

Under 6-BA treatment, ABA content in functional male flowers decreased then increased. ABA content in bisexual flowers was lower than that in functional male flowers at P1–P3 stages. During the bisexual flowers development, ABA content in the treatment was lower than that in the control at P1–P2 stages, but higher than that in the control at P3–P5 stages. During the functional male flowers development, ABA content in the treatment was higher than that in the control at P3–P4 stages (**Figure 9**).

Under 6-BA treatment, BR content in bisexual flowers showed a gradually increasing trend, the content of BR in functional male flowers showed the increasing- decreasing- increasing trend (**Figure 10**). Under 6-BA treatment, BR content in bisexual flowers was higher than that in functional male flowers at P1–P2 and P5–P7 stages, but lower at P3–P4 stages. At P1–P6 stages

of the functional male flowers development, the content of BR in the treatment was lower compared with that in the control.

After the induction of exogenous 6-BA, the content of JA exhibited a trend of firstly decreased, next increased, and then decreased during bisexual flowers development (**Figure 9**). Under 6-BA treatment, JA content in bisexual flowers was higher compared with functional male flowers at P1–P3 stages, whereas bisexual flowers were lower than functional male flowers at P4–P5 stages. At P1–P3 stage of the bisexual flowers development, the content of JA in the treatment was higher than that in the control.

Under 6-BA treatment, GA content in bisexual flowers increased then decreased, this change pattern was consistent with that in functional male flowers (**Figure 9**). Under 6-BA treatment, the concentration of GA in bisexual flowers was higher compared with functional male flowers (except P2). During P2–P7 stage for the bisexual flowers, a higher content of GA in the treatment was observed than that in the control. At P2 and P4–P7 stages for the functional male flowers, the content of GA in the treatment was higher than that in the control.

Expression of PgARF6 Genes in Response to IBA Treatment

To determine the response of *PgARF6* genes to exogenous IBA, the expression patterns of *PgARF6* in pomegranate flowers after IBA treatment were investigated using qRT-PCR. Under IBA treatment, the transcriptional level of *PgARF6a* in bisexual flowers was lower compared with that in functional male flowers at ovule development P2–P3 stages, and bisexual flowers was 7.7 times lower than functional male flowers at P3 (**Figure 10**). *PgARF6a* expression level of bisexual flowers was higher than functional male flowers at P4–P6. The transcriptional level of bisexual flowers was 2.8, 1.7, and 4.0 times higher than functional male flowers at P4, P5, and P6, respectively.

Induced by exogenous 6-BA, the expression level of *PgARF6b* in bisexual flowers was higher compared with that in functional male flowers during the whole development process (except P3). At the P3 stage, *PgARF6b* transcriptional level in bisexual flowers was 1.9 times lower than that in functional male flowers. At ovule development P2, P4, and P5 stages, *PgARF6b* expression level in bisexual flowers was 2.4, 1.32, and 3.9 times higher than that in functional male flowers, respectively.

Under 6-BA treatment, the transcriptional level of *PgARF6c* in bisexual flowers was higher than that in functional male flowers at the critical stage of ovule development (3.0–10.0 mm, P1–P3). At P4 stage of ovule sterility, *PgARF6c* expression level in bisexual flowers was 1.3 times lower than that in functional male flowers. At P5 stage, *PgARF6c* expression level in bisexual flowers was 2.4 times higher than functional male flowers.

Response of Endogenous Hormone in Pomegranate Flowers to Exogenous IBA Treatment

We further analyzed the content changes of endogenous hormone in pomegranate flower development under the treatment of exogenous IBA (**Figure 11**). Under IBA treatment, the content of IAA in bisexual flowers firstly increased, decreased and then increased during flower development. The concentration of IAA in bisexual flowers was lower than that in functional male flowers at P1–P2 and P4 stages, but higher than that in functional male flowers at P3, P5, and P7 stages. At the critical period of ovule development (P2–P5), the content of IAA in the treatment was higher compared with that in the control.

Under IBA treatment, ZR content in bisexual flowers firstly increased, decreased and then increased (**Figure 11**). Under IBA treatment, the concentration of ZR in bisexual flowers was lower

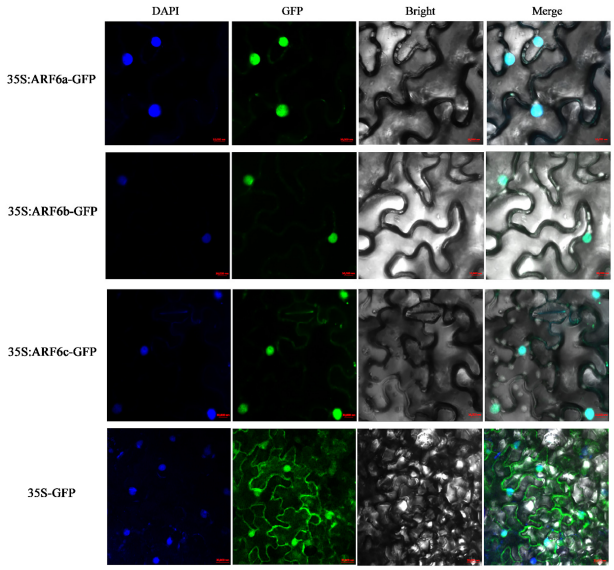


FIGURE 4 | Subcellular localization of PgARF6s-GFP fusion in tobacco.

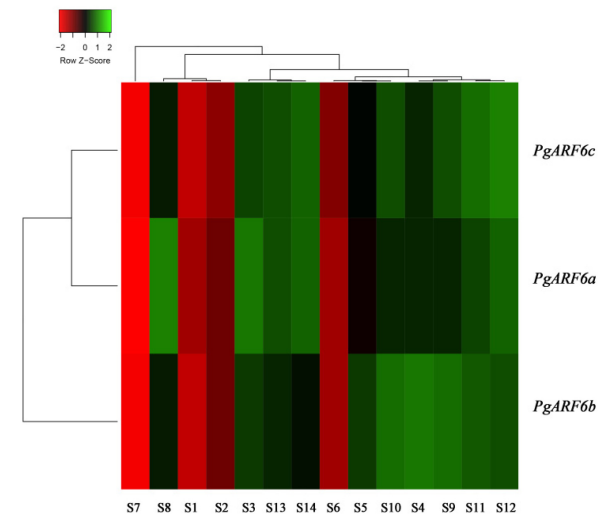


FIGURE 5 | *PgARF6* expression heatmap in different organs of pomegranate. S1, "Baiyushizi" inner seed coat; S2, "Tunisiruanzi" inner seed coat; S3, flower; S4, leaf; S5, root; S6, inner seed coat; S7, outer seed coat; S8, pericarp; S9, bisexual flower I; S10, functional male flower I; S11, bisexual flower II; S12, functional male flower II; S13, bisexual flower III; S14, functional male flower III (S3–S8 were "Dabenzi," S9–S14 were the pistil of "Tunisiruanzi").

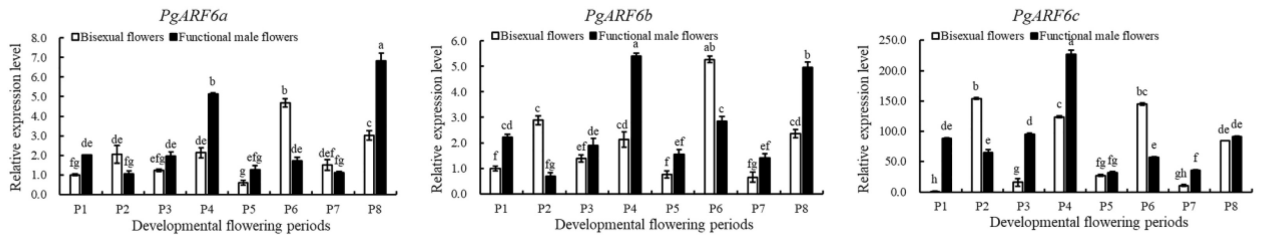


FIGURE 6 | Expression of *PgARF6s* transcripts during pomegranate flowers development.

TABLE 2 | Prediction of targeting relationship between *PgARF6s* and PgmiR167 in pomegranate.

miRNA	Target gene	Target start	Target end	MiRNA aligned fragment	Target aligned fragment
PgmiR167a	*PgARF6a*	2,370	2,390	UGAAGCUGCCAGCAUGAUCUA	GAGAUCAGGCUGGCAGCUUGU
PgmiR167a	*PgARF6b*	2,619	2,639	UGAAGCUGCCAGCAUGAUCUA	GAGAUCAGGCUGGCAGCUUGU
PgmiR167a	*PgARF6c*	2,382	2,402	UGAAGCUGCCAGCAUGAUCUA	GAGAUCAGGCUGGCAGCUUGU
PgmiR167d	*PgARF6b*	2,618	2,639	UGAAGCUGCCAGCAUGAUCUGG	AGAGAUCAGGCUGGCAGCUUGU
PgmiR167d	*PgARF6c*	2,381	2,402	UGAAGCUGCCAGCAUGAUCUGG	AGAGAUCAGGCUGGCAGCUUGU
PgmiR167d	*PgARF6a*	2,369	2,390	UGAAGCUGCCAGCAUGAUCUGG	UGAGAUCAGGCUGGCAGCUUGU

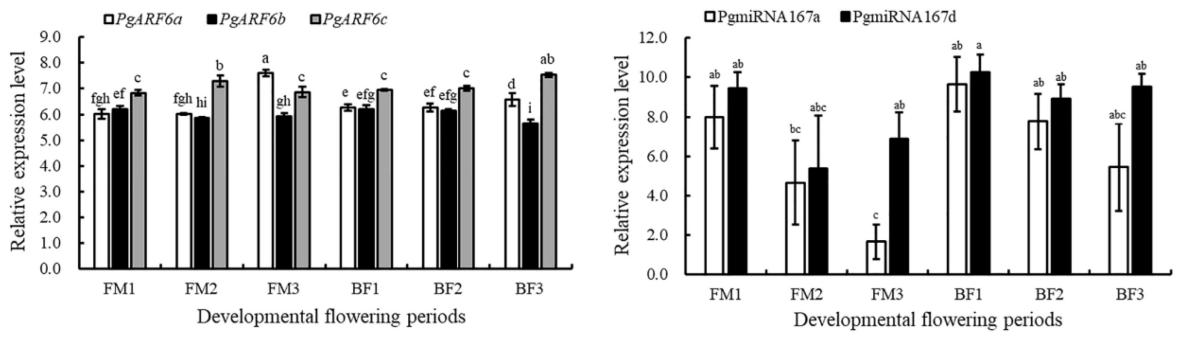


FIGURE 7 | Transcriptome and small-RNA sequencing results of *PgARF6*-PgmiR167 expression analysis.

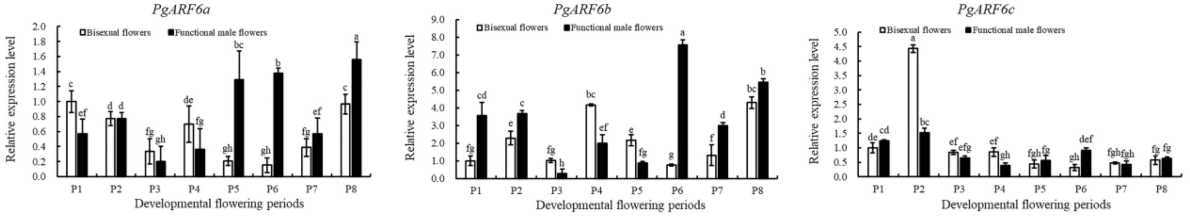


FIGURE 8 | Expression analyses of *PgARF6s* under 6-BA treatment.

than that in functional male flowers at P1–P2 and P4–P5 stages, but higher at P3 and P5 stages. During the bisexual flowers development, ZR content in the treatment was higher than that in the control at the critical stage of ovule development (P2–P4), but lower at P1 and P5 stages. During the functional male flowers development, ZR concentration in the treatment was higher than that in the control at P1–P3 stages, whereas at P4–P5 stages.

After external IBA induction, ABA content decreased then increased (P8) during bisexual flowers development (**Figure 11**). Under IBA treatment, the content of ABA in bisexual flowers was lower than that in functional male flowers at P2–P3 and P5–P8 stages, but higher at P1 and P4 stages. During the P2–P4 stages of functional male flowers ovule abortion, the concentration of endogenous ABA showed a decreasing trend.

During the development of the bisexual flowers, the content of BR under IBA treatment firstly increased, decreased and then increased (**Figure 11**). Under IBA treatment, the content of BR in bisexual flowers was lower compared with that in functional male flowers, except P8. At P1–P5 stages of bisexual flowers

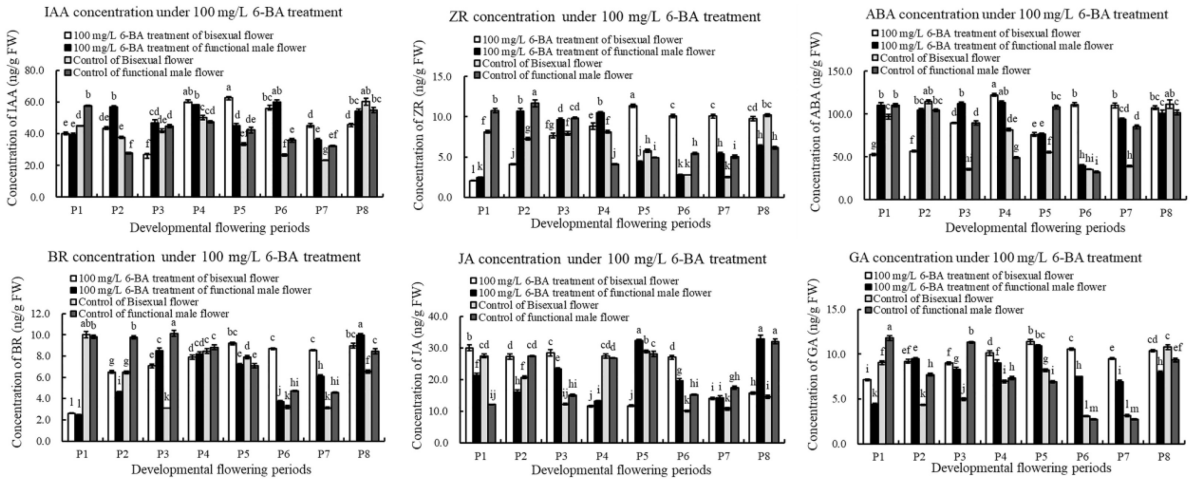


FIGURE 9 | Endogenous hormone in pomegranate flowers under 6-BA treatment.

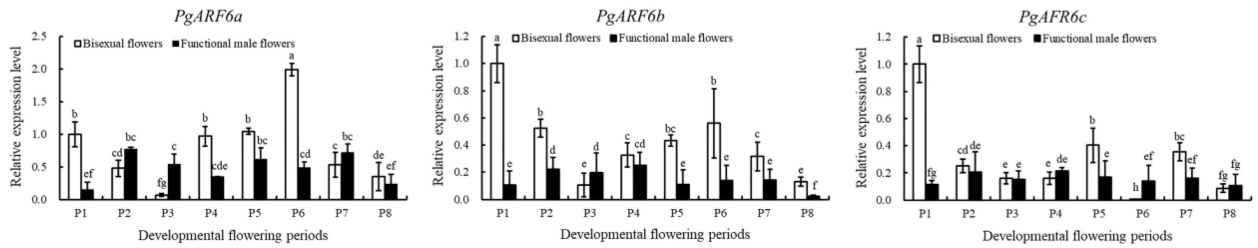


FIGURE 10 | Expression analyses of *PgARF6s* under IBA treatment.

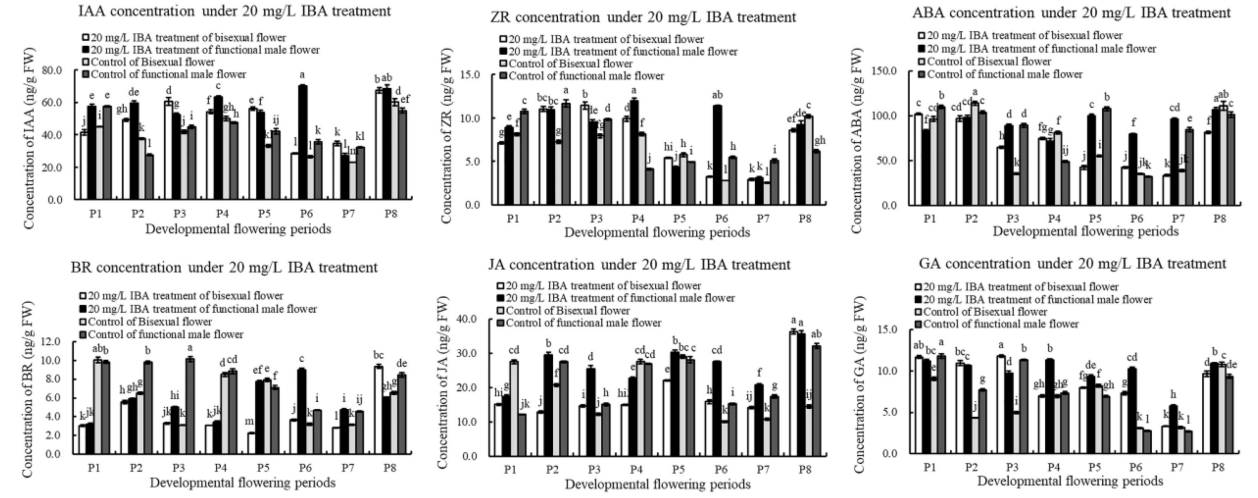


FIGURE 11 | Endogenous hormone in pomegranate flowers under IBA treatment.

ovule development, the concentration of BR in the treatment was lower than that in the control. During the functional male flowers development, BR content in the treatment was lower than that in the control at P1–P4 stages, but higher than that in the control at P5–P7 stages.

Under IBA treatment, the content of JA in bisexual flowers was lower compared with functional male flowers, except P8 (**Figure 11**). During the bisexual flowers development, JA content in the treatment was lower than that in the control at P1–P2 and P4–P5 stages, whereas BR content in the treatment was

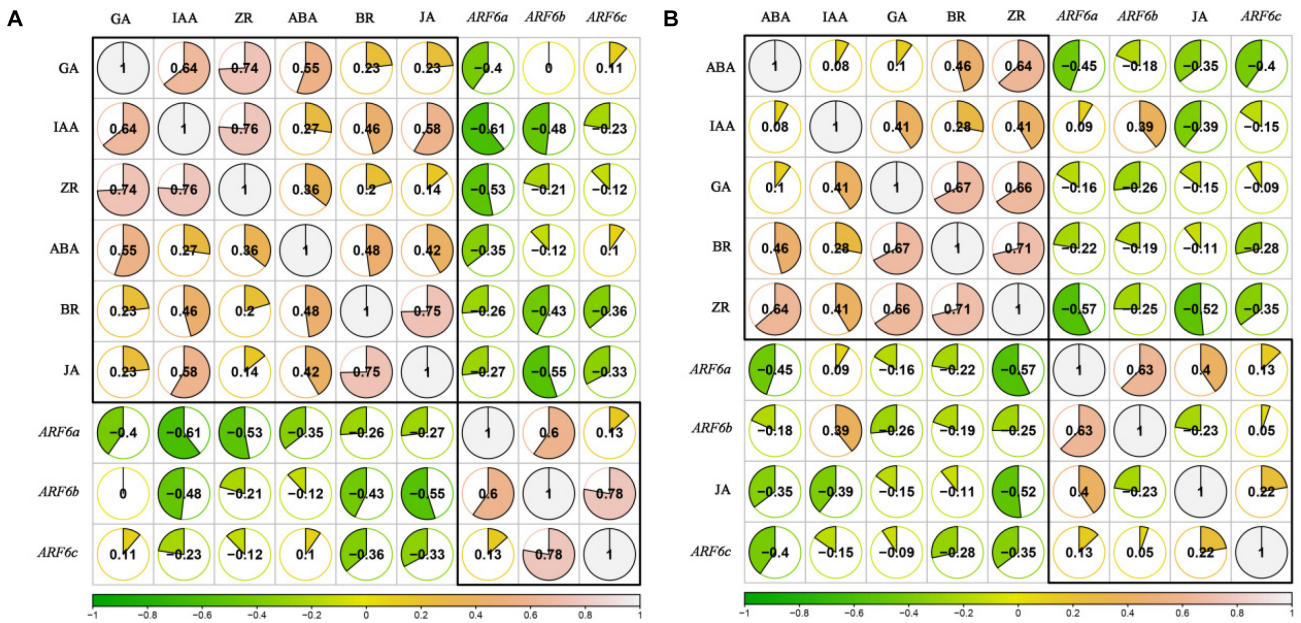


FIGURE 12 | (A) The correlation between endogenous hormones and *PgARF6* genes expression after IBA treatment. **(B)** The correlation between endogenous hormones and *PgARF6* genes expression after 6-BA treatment.

higher than that in the control at P3 and P6–P8 stages. During the functional male flowers development, JA concentration in the treatment was higher than that in the control, but lower at P4 stage.

Under exogenous IBA treatment, the content of GA in bisexual flowers was higher than that in functional male flowers at P1–P4 and P6–P7 stages, but lower at P5 and P8 (**Figure 11**). At the critical stage of ovule development in bisexual flower (3.0–12.0 mm, P1–P4), the concentration of GA in the treatment was higher compared with that in the control. During the functional male flowers development, GA content in treatment was higher than that in the control, except P1 and P3 stages.

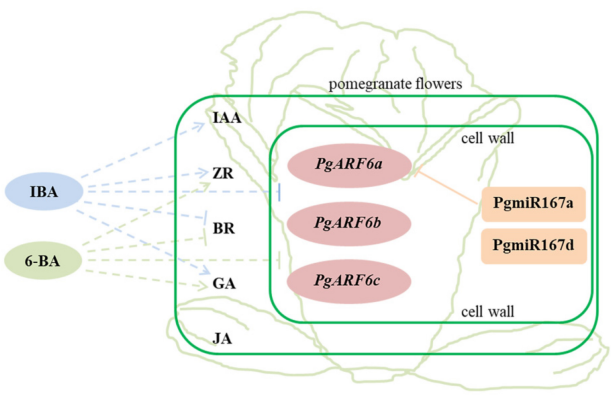


FIGURE 13 | The relationship model for hormones, PgARF6 genes and PgmiR167 in pomegranate flowers development. Solid line represents the direct relationship, dotted line represents the indirect relationship; arrow represents promotion and vertical line represents inhibition.

The Correlation Between Endogenous Hormones and PgARF6s Expression Levels After Exogenous IBA and 6-BA Treatment

In order to analyze the relationships between endogenous hormones content (IAA, ZR, BR, JA, GA, and ABA) and *PgARF6s* expression levels in pomegranate flowers development after exogenous IBA and 6-BA treatment, the correlation coefficient was calculated by R-rcorr and visualized by R-corrplot. As illustrated in **Figure 12A**. When treated with IBA, the endogenous BR, JA, IAA, and ZR levels were negatively correlated with the expressions of *PgARF6a*, *PgARF6b,* and *PgARF6c*. GA content was negatively correlated with *PgARF6a* expression, but positively correlated with the expression levels of *PgARF6b* and *PgARF6c*. ABA content was negatively correlated with the expression levels of *PgARF6a* and *PgARF6b*, but slightly positively correlated with *PgARF6c*. The pairwise positive correlation was observed between the endogenous IAA, ZR, BR, JA, GA, and ABA. There were the significantly positive correlation between GA, IAA, and ZR. JA was significantly positively correlated with BR. *PgARF6b* was significantly positively correlated with *PgARF6a* and *PgARF6c*, but the positive correlation between *PgARF6a* and *PgARF6c* was not significant.

Figure 12B showed the correlation between the content of endogenous hormones and the expression levels of *PgARF6s* after exogenous 6-AB treatment. The contents of endogenous BR, ZR, GA, and ABA were negatively correlated with the expression

levels of *PgARF6s*. IAA content was negatively correlated with *PgARF6c* expression level, but positively correlated with *ARF6a* and *PgARF6b* transcriptional levels. JA content was positively correlated with the expression levels *PgARF6a* and *PgARF6c*, but negatively correlated with *PgARF6b*. There were the positive correlation between the endogenous hormones ABA, IAA, ZR, GA, and BR. There was a significantly positive correlation between GA, BR, and ZR. ABA was significantly positive with ZR, and JA had a negative correlation with other endogenous hormones. *PgARF6a* was significantly positively correlated with *PgARF6b*, but *PgARF6c* was positively correlated with *PgARF6a* and *PgARF6b*. These results showed that endogenous hormones content and *PgARF6* genes expressions interact with each other, indicating a complex process required multiple factors to regulate pomegranate flower development.

DISCUSSION

PgARF6s Identification and Bioinformatics Analysis

ARF is involved in regulating multiple processes in plants and plays crucial roles in growth and development (Yao et al., 2019; Zheng et al., 2019; Mao et al., 2020). *ARFs* play an independent role in auxin signal transduction and participate in regulating embryogenesis, flowering and fruit-setting (Wu et al., 2011; Shi et al., 2014). Previous researches have shown that *ARFs* interact with *AG*, *AP2,* and miRNAs to participate in the morphogenesis of pistil (Liu N. et al., 2014; Liu X. G. et al., 2014). In our study, three *PgARF6* genes were identified and cloned from the "Tanshanhong" pomegranate. Bioinformatics analysis showed that there were numerous protein insertions and deletion sites in PgARF6b and PgARF6c sequence, and the tertiary structure of PgARF6 proteins were different, which might manifest function differentiation for *PgARF6s*. Protein interaction analysis showed that ARF6 directly interacted with AUX, IAA, and BEE2 to be involved in BR signal. The expression pattern of *PgARF6s* in pomegranate different organs showed that *PgARF6s* were involved in pomegranate flower development. Our results verified that PgARF6a, PgARF6b and PgARF6c proteins were localized in cell nucleus, which were consistent with the localization of CsARF and SlARF (Wu et al., 2011; Xu et al., 2016), indicating that PgARF6s were performed its functions in cell nucleus.

Expression Patterns of *PgARF6s* in Different Stages and Response to Hormone Treatments

Gene expression analysis from previous studies indicated that ARFs regulate numerous auxin-related processes at different plant developmental stages (Hu et al., 2015; Yu et al., 2020). Su et al.'s (2017) results suggested that *AcoARF3* and *AcoARF11*, the homologous gene of *AtARF6*, might be involved ovule development in pineapple fruit. The expression level of *ARF6* was higher than that of other *FvARFs* in each developing stage of strawberry fruit (Wang et al., 2019). *PgrARF1* and *PgrARF2* were highly expressed in both inner and outer pomegranate seed coat, inferring that these might involve in seed coat development through cell divisions in response to auxin regulation (Yu et al., 2020). The mRNA expression results demonstrated that the expression levels of *PgARF6s* in bisexual flowers were higher than in functional male flowers at the P2 stage of flower development. At the critical stages of ovule sterility P3–P5 and P7–P8 stages of flower organ maturation, *PgARF6s* expression levels in bisexual flowers were lower than that in functional male flowers, indicating that *PgARF6s* were related to ovule sterility and male flower organ maturation. Our results were consistent with the regulation of female sterility in tomatoes by ARF6-miR167 (Nagpal et al., 2005). In addition, we inferred that *PgARF6* might be a negative regulator of ovule development in pomegranate.

ARF6 and *ARF8* transcriptions are indeed suppressed by exogenously applied IBA (Zhang et al., 2017). In sweet orange, the expression of *CiARF6* genes was markedly down-regulated just after IAA treatment (Li et al., 2015). *PgARF6s* showed different expression patterns under the treatment of exogenous hormones. During the development of bisexual flowers, the expression levels of *PgARF6a* and *PgARF6c* in the 6-BA treatment were lower than that in the control at the critical stages of ovule development (5.1–14.0 mm, P2–P5). During the development of functional male flowers, the transcriptional levels of *PgARF6a* and *PgARF6c* in the 6-BA treatment were lower than that in the control at the critical stages of ovule development (5.1–12.0 mm, P2–P4). Under 6-BA treatment, *PgARF6b* transcriptional level in bisexual flowers was higher than that in functional male flowers at the P2 stage, whereas *PgARF6b* expression level in bisexual flowers were higher than that in functional male flowers at the P3–P5 stage. Under IBA treatment, the transcriptional level of *PgARF6b* in the treatment was lower than that in the control at the critical stages of ovule development (3.0–12.0 mm, P1–P4). During the development of bisexual flowers, the expression levels of *PgARF6a* and *PgARF6c* in the IBA treatment were lower compared with that in the control. During the development of functional male flowers, the expression levels of *PgARF6a* in the IBA treatment were lower than that in the control at the stages of ovule development (3.0–12.0 mm, P1–P4), and the expression levels of *PgARF6c* in the IBA treatment were lower compared with that in the control. Under the 6-BA treatment, *DnARF6* was weakly down-regulated in *D. officinale* seedlings, whereas *DnARF6* was significantly reduced by the IAA treatment (Chen Z. H. et al., 2017). These results indicated that the expressions of *PgARF6a*, *PgARF6b,* and *PgARF6c* were inhibited by the exogenous hormone treatment at the critical stages of ovule development (5.1–12.0 mm, P2–P4), indicating that the expressions of *PgARF6* genes were response to exogenous hormone treatments during the development of pomegranate flowers.

PgmiR167 Targets PgARF6s Expression in Pomegranate Flowers Development

In *Arabidopsis thaliana*, miR167 and target genes *ARF6* and *ARF8* regulate pistil and stamen development and maturity

(Wu et al., 2006). *ARF6* and *ARF8* are essential regulators for ovule and anther development. The flower organs of *arf6 arf8* double mutants show specific phenotypes, such as small petals, closed buds, and short stamens (Nagpal et al., 2005). MiR167 is involved in regulating ovule and stamen development by regulating the expression of *ARF6* and *ARF8* (Ru et al., 2006; Wu et al., 2006). In previous research, down-regulation of *ARF6* and *ARF8* by miR167 leads to floral development defects and female sterility in tomato (Liu N. et al., 2014). In this study, we found that PgmiR167 had only one target binding site on the mature mRNA sequence of *PgARF6s*. The *ARF8*, but not *ARF6*, mRNAs are degraded in the *35S-miR167b* transgenic plants (Ru et al., 2006). The analysis of transcriptome and small-RNA sequencings results showed that *PgARF6a* and PgmiR167a had a directly negative targeted regulation relationship. In addition, we also found that *PgARF6c* expression level was significantly higher than *PgARF6a* and *PgARF6b* in the floral organ development of pomegranate. However, there was no clear the targeted regulation relationship between *PgARF6c* and PgmiR167, suggesting that there may be specific functions for *PgARF6c*. Increased miR167 expression in the transgenic tomato plants lead to reduced transcriptional abundance of *ARF6* and *ARF8* (Liu N. et al., 2014). The results were similar to the previous reports (Wu et al., 2006; Yao et al., 2019), suggesting that the targeting relationship between miR167 and ARF6 was conserved in plants.

Response of Endogenous Hormone in Pomegranate Flowers to Exogenous Hormone Treatments

Plant hormones are essential signaling molecules that regulate various aspects of plant growth and development. Recent studies have indicated cytokinin in processes related to plant reproductive development and embryogenesis (Šimášková et al., 2015; Cucinotta et al., 2016; Hallmark and Rashotte, 2019). Auxin, BR, GA, and cytokinin are crucial for seed development, and numerous genes regulating seed size are correlated with BRs and auxin (Hu et al., 2018; Zhang X. F. et al., 2020). ABA and ethylene play roles in fruit development and maturation (Fenn and Giovannoni, 2021). Several recent studies suggest that interaction or balance of plant hormones regulate the development of flower buds (Sun et al., 2017; Zhao et al., 2021). Zu et al. (2021). The interaction of BR and cytokinin promotes ovule initiation and increases seed number per silique in *Arabidopsis* (Zu et al., 2021). Hormones are involved in plant development by affecting the expression levels of ARF genes in different plant species (Ha et al., 2013; Chen Z. H. et al., 2017). Our results showed that exogenous plant growth regulators affected the levels of endogenous hormones and *PgARF6* genes expression during pomegranate flower differentiation. *ARFs* in sweet orange and peach positively responded to exogenous auxin but display distinct expression patterns (Li et al., 2015; Diao et al., 2020). The expressions of *PgARF6* genes were inhibited under exogenous 6-BA and IBA treatments, which were consistent with the result of previous research (Chen Z. H. et al., 2017), indicating that *PgARF6s* expression were responsive to exogenous hormones IBA and 6-BA. At the critical stage of ovule sterility (5.1–14.0 mm, P2–P5), exogenous IBA promoted endogenous IAA and ZR accumulation, but inhibited BR. Both exogenous IBA and 6-BA promoted endogenous GA accumulation, and exogenous 6-BA promoted endogenous IAA accumulation at the critical stage of pomegranate ovule development. In the flower bud differentiation stage of "Mantianhong" × "Dangshansu pear" hybrid offspring, the contents of ZT and ABA were increased, but the contents of GA_3 and IAA were decreased by spraying 100 mg/kg 6-BA (Liu et al., 2015). After 6-BA treatment, the content of IAA in the physiological differentiation stage of apple flower buds was significantly lower than that in the control (Li et al., 2015). After IAA treatment, the contents of endogenous IAA and GAs significantly increased during the physiological differentiation of "Fuji" apple flower buds (Li, 2015). This research was consistent with our result, indicating that exogenous auxin promoted endogenous IAA and cytokinin accumulation in floral organ of plant.

PgARF6s Expression Coordinated Endogenous Hormones Response to Exogenous Hormone Treatments in Pomegranate Flower Development

The results of correlation analysis showed that the content of endogenous hormones was negatively correlated with the expression levels of *PgARF6s* during the development of pomegranate flowers. Under exogenous 6-BA treatment, JA and *PgARF6* genes were grouped together, and JA was negatively correlated with ZR, BR, ABA, and GA. JA content was positively correlated with the expression levels of *PgARF6a* and *PgARF6c*, indicating a significant correlation between the content of JA and *PgARF6s* expressions in the pomegranate flower development. The above results indicated that the interaction between endogenous hormones and genes expressions was jointly involved in regulating the development of pomegranate flower.

In our study, three *PgARF6* genes have the same gene structure and conserved domains. *PgARF6s* expression levels in bisexual flowers were significantly lower than those in functional male flowers at the critical stage of ovule sterility (8.1–14.0 mm), indicating that *PgARF6* genes had the conservative function of involving in ovule sterility developing into functional male flower. However, numerous differences were found in PgARF6 proteins sequences and tertiary structure. *PgARF6s* expression patterns were different in pomegranate flower and under exogenous hormones treatment, the expression of *PgARF6c* was mightily inhibited by the exogenous hormone, and *PgARF6a* was directly regulated by PgmiR167a, indicating that *PgARF6* genes performed different function in pomegranate flower development. Based on experimental data, we proposed a relationship model for hormones, *PgARF6* genes and PgmiR167 in pomegranate flowers (**Figure 13**). PgARF6s functioned in cell nucleus, and PgmiR167a directly targeted *PgARF6a* and affected its expression. Under hormone treatment, the expressions of three *PgARF6* genes were inhibited by exogenous IBA and 6-BA. At the

critical stage of ovule sterility (8.1–12.0 mm), exogenous IBA promoted the accumulation of IAA, GA, and ZR, and reduced the accumulation of BR. When the vertical diameter was 8.1–12.0 mm, exogenous 6-BA promoted the content of ZR and GA accumulation, and inhibited BR accumulation. Under 6-BA treatment, ZR accumulation was negatively correlated with the expressions of *PgARF6* genes. Under IBA treatment, the content of BR, IAA, and JA were negatively correlated with the expressions of *PgARF6* genes. Our results suggested that *PgARF6* genes were involved in ovule sterility of pomegranate flowers.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

YZ and ZY conceived and planned all experiments. YZ analyzed and cloned *PgARF6* subfamily gene members and wrote the manuscript. YZ and YW performed the expression level examination. YZ and YR helped to prepare datasets and figures for the manuscript. YZ and MY were the major contributors in analyzing data. XZ and ZY revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Initiative Project for Talents of Nanjing Forestry University (GXL2014070), the Priority Academic Program Development of Jiangsu High Education Institutions (PAPD), China Scholarship Council (CSC, 202008320482), and the National Natural Science Foundation of China (31901341). These funding bodies took part in the design of the study and collection, analysis, interpretation of data, and the writing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.833747/full#supplementary-material

REFERENCES

Abel, S., and Theologis, A. (1996). Early genes and auxin action. *Plant Physiol.* 111, 9–17. doi: 10.1104/pp.111.1.9

Cai, Y. L., Lu, X. G., and Zhu, L. W. (1993). Preliminary research on flower bud differentiation of 'Pink' pomegranate. *Yuan Yi Xue Bao* 20, 23–26.

Chen, L. N., Niu, J., Liu, B. B., Jing, D., Lou, X., Li, H. X., et al. (2020). Effects of foliar application of different plant growth regulators on fruit setting rate and quality in pomegranate. *J. Fruit Sci.* 37, 244–253.

Chen, L. N., Zhang, J., Li, H. X., Niu, J., Xue, H., Liu, B. B., et al. (2017). Transcriptomic analysis reveals candidate genes for female sterility in pomegranate flowers. *Front. Plant Sci.* 8:1430. doi: 10.3389/fpls.2017.01430

Chen, Z. H., Yuan, Y., Fu, D., Shen, C. J., and Yang, Y. J. (2017). Identification and expression profiling of the Auxin response factors in *Dendrobium officinale* under abiotic stresses. *Int. J. Mol. Sci.* 18:927. doi: 10.3390/ijms18050927

Cucinotta, M., Manrique, S., Guazzotti, A., Quadrelli, N. E., Mendes, M. A., Benkova, E., et al. (2016). Cytokinin response factors integrate auxin and cytokinin pathways for female reproductive organ development. *Development* 143, 4419–4424. doi: 10.1242/dev.143545

Damodharan, S., Zhao, D. Z., and Arazi, T. (2016). A common miRNA160-based mechanism regulates ovary patterning, floral organ abscission and lamina outgrowth in tomato. *Plant J.* 86, 458–471. doi: 10.1111/tpj.13127

Diao, D. H., Hu, X., Guan, D., Wang, W., Yang, H. Q., and Liu, Y. P. (2020). Genome−wide identifcation of the ARF (auxin response factor) gene family in peach and their expression analysis. *Mol. Biol. Rep.* 47, 4331–4344. doi: 10.1007/s11033-020-05525-0

Ellis, C. M., Nagpal, P., Young, J. C., Hagen, G., Guilfoyle, T. J., and Reed, J. W. (2005). AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*. *Development* 132, 4563–4574. doi: 10.1242/dev.02012

Feng, Z., Zhu, J., Du, X., and Cui, X. (2012). Effects of three auxin-inducible LBD members on lateral root formation in *Arabidopsis thaliana*. *Planta* 236, 1227–1237. doi: 10.1007/s00425-012-1673-3

Fenn, M. A., and Giovannoni, J. J. (2021). Phytohormones in fruit development and maturation. *Plant J.* 105, 446–458.

Fukaki, H., Nakao, Y., Okushima, Y., Theologis, A., and Tashaka, M. (2005). Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in *Arabidopsis*. *Plant J.* 44, 382–395. doi: 10.1111/j.1365-313X.2005.02537.x

Guilfoyle, T. J., and Hagen, G. (2007). Auxin response factors. *Curr. Opin. Plant Biol.* 10, 453–460.

Guilfoyle, T. J., Ulmasov, T., and Hagen, G. (1998). The ARF family of transcription factors and their role in plant hormone-responsive transcription. *Cell. Mol. Life Sci.* 54, 619–627. doi: 10.1007/s000180050190

Ha, C. V., Le, D. T., Nishiyama, R., Watanabe, Y., Sulieman, S., Tran, U. T., et al. (2013). The auxin response factor transcription factor family in soybean: genome-wide identification and expression analyses during development and water stress. *DNA Res.* 20, 511–524. doi: 10.1093/dnares/dst027

Hallmark, H. T., and Rashotte, A. M. (2019). Cytokinin response factors: responding to more than cytokinin. *Plant Sci.* 289:110251. doi: 10.1016/j.plantsci.2019.110251

Hardtke, C. S., and Berleth, T. (1998). The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO J.* 17, 1405–1411. doi: 10.1093/emboj/17.5.1405

Hu, W., Zuo, J., Hou, X., Yan, Y., Wei, Y., Liu, J., et al. (2015). The auxin response factor gene family in banana: genome-wide identification and expression analyses during development, ripening, and abiotic stress. *Front. Plant Sci.* 6:742. doi: 10.3389/fpls.2015.00742

Hu, Z., Lu, S. J., Wang, M. J., He, H., Sun, L., Wang, H., et al. (2018). A novel QTL qTGW3 encodes the GSK3/SHAGGY-like kinase OsGSK5/OsSK41 that interacts with OsARF4 to negatively regulate grain size and weight in rice. *Mol. Plant* 11, 736–749. doi: 10.1016/j.molp.2018.03.005

Huang, X. B., Zhang, T. K., Liu, C. Y., Zhao, Y. J., Wei, H. M., Zhou, J. Q., et al. (2019). Genome- wide identification and expression analysis of auxin response factor (ARF) gene family in *Punica granatum*. *J. Fruit Sci.* 36, 43–55.

Hussain, S., Kim, S. H., Bahk, S., Ali, A., Canh, N. X., Yun, D. J., et al. (2020). The auxin signaling repressor IAA8 promotes seed germination through down-regulation of ABI3 transcription in *Arabidopsis*. *Front. Plant Sci.* 11:111. doi: 10.3389/fpls.2020.00111

Johanningsmeier, S. D., and Harris, G. K. (2011). Pomegranate as a functional food and nutraceutical source. *Annu. Rev. Food Sci. Technol.* 2, 181–201. doi: 10.1146/annurev-food-030810-153709

Johnson, L. A., and Douglas, C. J. (2007). Populus trichocarpa monopteros/auxin response factor5 (ARF5) genes: comparative structure, sub-functionalization,

and Populus-Arabidopsis microsynteny. *Can. J. Bot.* 85, 1058–1070. doi: 10.1139/b07-065

Kumar, R., Tyagi, A. K., and Sharma, A. K. (2011). Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. *Mol. Genet. Genomics* 285, 245–260. doi: 10.1007/s00438-011-0602-7

Li, H. F., Zhang, W. Q., Dong, Q. L., Wang, X. F., and Ran, K. (2018). Cloning, sequencing and expression analysis of auxin response factors (*MdARF*) in apple. *J. Fruit Sci.* 35, 1170–1181.

Li, J., Dai, X., and Zhao, Y. (2006). A role for auxin response factor 19 in auxin and ethylene signaling in *Arabidopsis*. *Plant Physiol.* 140, 899–908. doi: 10.1104/pp.105.070987

Li, Q. F., Lu, J., Yu, J. W., Zhang, C. Q., and He, J. X. (2018). The brassinosteroid-regulated transcription factors BZR1/BES1 function as a coordinator in multisignal-regulated plant growth. *Biochim. Biophys. Acta Gene Regul. Mech.* 6, 561–571. doi: 10.1016/j.bbagrm.2018.04.003

Li, S. B., Ouyang, W. Z., Hou, X. J., Xie, L. L., Hu, C. G., and Zhang, J. Z. (2015). Genome-wide identification, isolation and expression analysis of auxin response factor (ARF) gene family in sweet orange (*Citrus sinensis*). *Front Plant Sci.* 6:119. doi: 10.3389/fpls.2015.00119

Li, Y. M. (2015). Effect of 6-BA and IAA Spraying On Hormone And Related Gene Expression During Floral Induction And Initiation in 'Fuji' Apple (Malus Domestica Borkh). Yangling: Northwest A&F University.

Liu, L., Wang, C., Yao, G. F., Wang, D. F., and Wu, J. (2015). Effects of exogenous growth regulator treatment on floral initiation of pear progenies from hybrids 'Mantianhong'×'Dangshansuli'. *J. Nanjing Agric. Univ.* 38, 381–388.

Liu, N., Wu, S., Houten, J. V., Wang, Y., Ding, B., Fei, Z. J., et al. (2014). Down-regulation of *AUXIN RESPONSE FACTORS 6* and *8* by microRNA 167 leads to floral development defects and female sterility in tomato. *J. Exp. Bot.* 65, 2507–2520. doi: 10.1093/jxb/eru141

Liu, X. G., Dinh, T. T., Li, D. M., Shi, B. H., Li, Y. P., Cao, X. W., et al. (2014). AUXIN RESPONSE FACTOR 3 integrates the functions of *AGAMOUS* and *APETALA2* in floral meristem determinacy. *Plant J.* 80, 629–641. doi: 10.1111/tpj.12658

Mao, Z. L., He, S. B., Xu, F., Wei, X. X., Jiang, L., Liu, Y., et al. (2020). Photoexcited CRY1 and phyB interact directly with ARF6 and ARF8 to regulate their DNA-binding activity and auxin-induced hypocotyl elongation in *Arabidopsis*. *New Phytol.* 225, 848–865. doi: 10.1111/nph.16194

Miyashima, S., Sebastian, J., Lee, J. Y., and Helariutta, A. (2013). Stem cell function during plant vascular development. *EMBO J.* 32, 178–193.

Nagpal, P., Ellis, C. M., Weber, H., Ploense, S. E., and Barkaw, L. S. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132, 4107–4118. doi: 10.1242/dev.01955

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tashaka, M. (2007). ARF7 and *ARF19* regulate lateral root formation via direct activation of LBD/ASL genes in *Arabidopsis*. *Plant Cell* 19, 118–130. doi: 10.1105/tpc.106.047761

Pekker, I., Alvarez, J. P., and Eshed, Y. (2005). Auxin response factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity. *Plant Cell* 17, 2899–2910. doi: 10.1105/tpc.105.034876

Piya, S., Shrestha, S. K., Binder, B., Stewart, Jr, C.N, and Hewezi, T. (2014). Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in *Arabidopsis*. *Front. Plant Sci.* 5:744. doi: 10.3389/fpls.2014.00744

Qin, G. H., Xu, C. Y., Ming, R., Tang, H. B., Guyot, R., Kramer, E. M., et al. (2017). The pomegranate (*Punica granatum* L.) genome and the genomics of punicalagin biosynthesis. *Plant J.* 91, 1108–1128. doi: 10.1111/tpj.13625

Roosjen, M., Paque, S., and Weijers, D. (2017). Auxin response factors: output control in auxin biology. *J. Exp. Bot.* 69, 179–188. doi: 10.1093/jxb/erx237

Ru, P., Xu, L., Ma, H., and Huang, H. (2006). Plant fertility defects induced by the enhanced expression of microRNA167. *Cell Res.* 16, 457–465. doi: 10.1038/sj.cr.7310057

Sessions, A., Nemhauser, J. L., Mccall, A., Roe, J. L., and Feldmann, K. A. (1997). Ettin patterns the *Arabidopsis* floral meristem and reproductive organs. *Development* 124, 4481–4491. doi: 10.1242/dev.124.22.4481

Shi, M. Y., Zhang, W., Yu, J., Wang, W. P., and Liu, Y. P. (2014). Cloning and expression analysis of *ARF* and *Aux/IAA* gene family members in Peach. *Acta Hortic. Sin.* 41, 536–544.

Šimášková, M., O'Brien, J. A., Khan, M., Noorden, G. V., Ötvös, K., Vieten, A., et al. (2015). Cytokinin response factors regulate PIN-FORMED auxin transporters. *Nat. Commun.* 6:8717. doi: 10.1038/ncomms9717

Su, Z. X., Wang, L. L., Li, W. M., Zhao, L. H., Huang, X. Y., Azam, S. M., et al. (2017). Genome-wide identification of auxin response factor (ARF) genes family and its tissue-specific prominent expression in pineapple (*Ananas comosus*). *Tropical Plant Biol.* 10, 86–96.

Sun, P., Li, J. R., Du, G. G., Han, W. J., Fu, J. M., Diao, S. F., et al. (2017). Endogenous phytohormone profiles in male and female floral buds of the persimmons (*Diospyros kaki* Thunb.) during development. *Sci. Hortic.* 218, 213–221.

Tang, Y. W. (2016). *Functional Studies Of Solanum Lycopersicum Auxin Response Factor 6* (*SlARF6) in Fruit Initiation And Development*. Chong Qing Shi: Chongqing University.

Teixeira da Silva, J. A., Rana, T. S., Narzary, D., Verma, N., Meshram, D. T., and Ranade, S. A. (2013). Pomegranate biology and biotechnology: a review. *Sci. Hortic.* 160, 85–107.

Tian, C. E., Muto, H., Higuchi, K., Matamura, T., Tatematsu, K., Koshiba, T., et al. (2004). Disruption and overexpression of auxin response factor 8 gene of *Arabidopsis* affect hypocotyl elongation and root growth habit,indicating its possible involvement in auxin homeostasis in light condition. *Plant J.* 40, 333–343. doi: 10.1111/j.1365-313X.2004.02220.x

Tian, H. Y., Lv, B. S., Ding, T. T., Bai, M. Y., and Ding, Z. J. (2018). Auxin-BR interaction regulates plant growth and development. *Front. Plant Sci.* 8:2256. doi: 10.3389/fpls.2017.02256

Tiwari, S. B., Hagen, G., and Guilfoyle, T. (2003). The roles of auxin response factor domains in auxin- responsive transcription. *Plant Cell* 15, 533–543. doi: 10.1105/tpc.008417

Ulmasov, T., Hagen, G., and Guilfoyle, T. J. (1997). *ARF1*, a transcription factor that binds to auxin response elements. *Science* 276, 1865–1868. doi: 10.1126/science.276.5320.1865

Wan, S. B., Li, W. L., Zhu, Y. Y., Liu, Z. M., Huang, W. D., and Zhan, J. C. (2014). Genome-wide identification,characterization and expression analysis of the auxin response factor gene family in *Vitis vinifera*. *Plant Cell Rep.* 33, 1365–1375. doi: 10.1007/s00299-014-1622-7

Wang, S. X. (2019). *Bioinformatics Analysis Of Auxin Response Factor* (*ARF*) Gene Family And Functional Research Of MiR167 And Its Target *ARF6* In Strawberry. Shenyang: Shenyang Agricultural University.

Wang, S. X., Shi, F. Y., Dong, X. X., Li, Y. X., Zhang, Z. H., and Li, H. (2019). Genome-wide identification and expression analysis of auxin response factor (*ARF*) gene family in strawberry (*Fragaria vesca*). *J. Integr. Agric.* 18, 1587–1603. doi: 10.1016/s2095-3119(19)62556-6

Wu, J., Wang, F. Y., Cheng, L., Kong, F. L., Peng, Z., Liu, S. Y., et al. (2011). Identification, isolation and expression analysis of auxin response factor (*ARF*) genes in *Solanum lycopersicum*. *Plant Cell Rep.* 30, 2059–2073. doi: 10.1007/s00299-011-1113-z

Wu, M. F., Tian, Q., and Reed, J. W. (2006). *Arabidopsis* microRNA167 controls patterns of *ARF6* and *ARF8* expression and regulates both female and male reproduction. *Development* 133, 4211–4218. doi: 10.1242/dev.02602

Xu, Y. X., Mao, J., Chen, W., Qian, T. T., Liu, S. C., Hao, W. J., et al. (2016). Identification and expression profiling of the auxin response factors (ARFs) in the tea plant [*Camellia sinensis (L.) O. Kuntze*] under various abiotic stresses. *Plant Physiol. Biochem.* 98, 46–56. doi: 10.1016/j.plaphy.2015.11.014

Yang, J., Tian, L., Sun, M. X., Huang, X. Y., Zhu, J., Guan, Y. F., et al. (2013). AUXIN RESPONSE FACTOR17 is essential for pollen wall pattern formation in *Arabidopsis*. *Plant Physiol.* 162, 720–731. doi: 10.1104/pp.113.214940

Yao, X. Z., Chen, J. L., Zhou, J., Yu, H. C. Z., Ge, C. N., Zhang, M., et al. (2019). An essential role for miRNA167 in maternal control of embryonic and seed development. *Plant Physiol.* 180, 453–464. doi: 10.1104/pp.19.00127

Yu, L. A., Liu, C. Y., Li, J. Y., Jia, B. T., Qi, X. X., Ming, R., et al. (2020). Identification of candidate Auxin response factors involved in pomegranate seed coat development. *Front. Plant Sci.* 11:536530. doi: 10.3389/fpls.2020.536530

Yuan, Z. H. (2015). *Chinese Fruit Tree Science and Practice-Pomegranate*. Xi'an: Shaanxi Science and Technology Press.

Yuan, Z. H., Fang, Y. M., Zhang, T. K., Fei, Z. J., Han, F. M., Liu, C. Y., et al. (2018). The pomegranate (*Punica granatum* L.) genome provides insights into fruit

quality and ovule developmental biology. *Plant Biotechnol. J.* 16, 1363–1374. doi: 10.1111/pbi.12875

Zhang, F., Tao, W. Q., Sun, R. Q., Wang, J. X., Li, C. L., Kong, X. P., et al. (2020). PRH1 mediates ARF7-LBD dependent auxin signaling to regulate lateral root development in *Arabidopsis thaliana*. *PLoS Genet.* 16:e1008044. doi: 10.1371/journal.pgen.1008044

Zhang, G. Z., Jin, S. H., Li, P., Jiang, X. Y., Li, Y. J., and Hou, B. K. (2017). Ectopic expression of *UGT84A2* delayed flowering by indole-3-butyric acid-mediated transcriptional repression of *ARF6* and *ARF*8 genes in Arabidopsis. *Plant Cell Rep.* 36, 1995–2006. doi: 10.1007/s00299-017-2225-x

Zhang, X. F., Tong, J. H., Bai, A. N., Liu, C. M., Xiao, L. T., and Xue, H. W. (2020). Phytohormone dynamics in developing endosperm influence rice grain shape and quality. *J. Integr. Plant Biol.* 62, 1625–1637. doi: 10.1111/jipb.12927

Zhao, Y. J., Liu, C. Y., Ge, D. P., Yan, M., Ren, Y., Huang, X. B., et al. (2020). Genome-wide identification and expression of YABBY genes family during flower development in *Punica granatum* L. *Gene* 752:144784. doi: 10.1016/j.gene.2020.144784

Zhao, Y. J., Liu, C. Y., Yan, M., Zhao, X. Q., Baloch, A. M., Baloch, A. W., et al. (2021). Variation of endogenous phytohormone in functional male and bisexual flowers of pomegranate (*Punica granatum* L.) during Development. *Pakistan J. Bot.* 53, 551–558. doi: 10.30848/PJB2021-2(12

Zhao, Y., Weng, Q. Y., Ma, H. L., Song, J. H., Yuan, J. C., Wang, L. Y., et al. (2016). Genome-wide identification and bioinformatics analysis of ARF gene family in foxtail millet *Setaria italica*. *J. Plant Genet. Resour.* 17, 547–554.

Zheng, L. J., Nagpal, P., Villarino, G., Trinidad, B., Bird, L., Huang, Y., et al. (2019). MiR167 limits anther growth to potentiate anther dehiscence. *Development* 146:174375. doi: 10.1242/dev.174375

Zu, S. H., Jiang, Y. T., Chang, J. H., Zhang, Y. J., Xue, H. W., and Lin, W. H. (2021). Interaction of brassinosteroid and cytokinin promotes ovule initiation and increases seed number per silique in *Arabidopsis*. *J. Integr. Plant Biol.* 1–39. doi: 10.1111/jipb.13197

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 25 March 2022
doi: 10.3389/fpls.2022.820439

Genome-Wide Identification and Expression Analysis of the R2R3-MYB Transcription Factor Family Revealed Their Potential Roles in the Flowering Process in Longan (*Dimocarpus longan*)

Qinchang Chen[1,2], Xiaodan Zhang[3], Yaxue Fang[4], Baiyu Wang[4], Shaosi Xu[1,2], Kai Zhao[1,2], Jisen Zhang[4] and Jingping Fang[1,2]**

[1] College of Life Sciences, Fujian Normal University, Fuzhou, China, [2] Center of Engineering Technology Research for Microalgae Germplasm Improvement of Fujian, Southern Institute of Oceanography, Fujian Normal University, Fuzhou, China, [3] Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, United States, [4] Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Key Laboratory of Genetics, Breeding and Multiple Utilization of Crops, Ministry of Education, Fujian Agriculture and Forestry University, Fuzhou, China

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Fanchang Zeng,
Shandong Agricultural University, China
Ziliang Luo,
University of Florida, United States

***Correspondence:**
Jingping Fang
jinphia@163.com
Jisen Zhang
zjisen@fafu.edu.cn

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

Received: 23 November 2021
Accepted: 02 March 2022
Published: 25 March 2022

Citation:
Chen Q, Zhang X, Fang Y, Wang B, Xu S, Zhao K, Zhang J and Fang J (2022) Genome-Wide Identification and Expression Analysis of the R2R3-MYB Transcription Factor Family Revealed Their Potential Roles in the Flowering Process in Longan (Dimocarpus longan). Front. Plant Sci. 13:820439. doi: 10.3389/fpls.2022.820439

Longan (*Dimocarpus longan* Lour.) is a productive fruit crop with high nutritional and medical value in tropical and subtropical regions. The *MYB* gene family is one of the most widespread plant transcription factor (TF) families participating in the flowering regulation. However, little is known about the MYB TFs involved in the flowering process in longan and its regulatory network. In this study, a total of 119 *DlR2R3-MYB* genes were identified in the longan genome and were phylogenetically grouped into 28 subgroups. The groupings were supported by highly conserved gene structures and motif composition of *DlR2R3-MYB* genes in each subgroup. Collinearity analysis demonstrated that segmental replications played a more crucial role in the expansion of the *DlR2R3-MYB* gene family compared to tandem duplications, and all tandem/segmental duplication gene pairs have evolved under purifying selection. Interspecies synteny analysis among longan and five representative species implied the occurrence of gene duplication events was one of the reasons contributing to functional differentiation among species. RNA-seq data from various tissues showed *DlR2R3-MYB* genes displayed tissue-preferential expression patterns. The pathway of flower development was enriched with six *DlR2R3-MYB* genes. *Cis*-acting element prediction revealed the putative functions of *DlR2R3-MYB* genes were related to the plant development, phytohormones, and environmental stresses. Notably, the orthologous counterparts between Arabidopsis and longan R2R3-MYB members tended to play conserved roles in the flowering regulation and stress responses. Transcriptome profiling on off-season flower induction (FI) by $KClO_3$ indicated two up-regulated and four

down-regulated *DlR2R3-MYB* genes involved in the response to $KClO_3$ treatment compared with control groups. Additionally, qRT-PCR confirmed certain genes exhibited high expression in flowers/flower buds. Subcellular localization experiments revealed that three predicted flowering-associated MYB proteins were localized in the nucleus. Future functional studies on these potential candidate genes involved in the flowering development could further the understanding of the flowering regulation mechanism.

Keywords: *Dimocarpus longan*, *R2R3-MYB*, flower development, $KClO_3$, qRT-PCR

INTRODUCTION

Myeloblastosis (MYB) proteins were first identified in the avian myeloblastosis virus (AMV) (Klempnauer et al., 1982), but have since been found in all eukaryote genomes and belong to one of the largest transcription factor (TF) gene families in plants. The first plant *MYB* gene identified from *Zea mays,* known as *COLORED1* (*C1*), was involved in the regulation of anthocyanin biosynthesis (Paz-Ares et al., 1987). The MYB TF contains a highly conserved MYB DNA-binding domain (DBD) in the N-terminal region that is usually composed of 1–4 serial imperfect repeat sequences. Each repeat sequence covers about 50–55 amino acid residues including three regularly spaced tryptophan (W) residues [-W-(X_{19})-W-(X_{19})-W. . .-F/I-(X_{-18})-W–(X_{-18})-W-], which fold into three α-helices (Kanei-Ishii et al., 1990; Kranz et al., 2000). The second and third helices form a helix-turn-helix (HTH) structure, and the third α-helix directly participates in binding with DNA (Ogata et al., 1992; Williams and Grotewold, 1997; Jia et al., 2004). MYB TFs can be divided into four categories based on the number of MYB repeats in their DBDs: 4R-MYB (four repeats), R1R2R3-MYB (three repeats), R2R3-MYB (two repeats), and 1RMYB (one repeat). The *R2R3-MYB* genes have received significant attention because they are massively expanded in the plant lineage and involved in regulating a wide array of plant-specific biological processes, such as primary and secondary metabolism, response to plant hormones and environmental stresses, disease resistance and leaf morphogenesis (Roy, 2016; Millard et al., 2019a; Jiang and Rao, 2020). The evolutionary history of the *R2R3-MYB* (*2R-MYBs*) gene family throughout the eukaryotic kingdom remains unknown (Du et al., 2015). The "loss" model proposes that *R2R3-MYB* genes potentially evolved from *R1R2R3-MYB* progenitor genes through the loss of the R1 repeat sequence (Rosinski and Atchley, 1998; Braun and Grotewold, 1999). In contrast, the more convincing "gain" model hypothesizes that *R1R2R3-MYB* genes evolved from the *R2R3-MYB*s via intragenic structural domain duplication (Jiang et al., 2004; Du et al., 2015).

With the rapid development of high-throughput sequencing technologies, genome-wide identification of R2R3-MYB proteins has been achieved in many plant species, for instance, *Arabidopsis thaliana* (Stracke et al., 2001; Dubos et al., 2010), *Medicago truncatula* (Li et al., 2019), *Casuarina equisetifolia* (Wang et al., 2021), *Solanum tuberosum* (Sun et al., 2019), *Oryza sativa* (Katiyar et al., 2012), and *Capsicum annuum* (Wang et al., 2020). In general, the number of R2R3-MYBs in plants ranged in size from 70 to 285, and these *MYB* genes evolved mainly based on the natural evolution of organisms and the recombination and amplification of genomes (Chang and Duda, 2012). The highest number of R2R3-MYBs (285) was found in the banana genome (Pucker et al., 2020). Respectively, 244, 192, 157, and 134 *R2R3-MYB* genes have been identified in the soybean (Du et al., 2012b), populus (Wilkins et al., 2009), maize (Du et al., 2012a), and grape (Wong et al., 2016) based on previous studies. In addition, the lowest number of *R2R3-MYB* genes, a total of 70, were identified in the sugar beet genome (Stracke et al., 2014).

R2R3-MYB TFs as regulatory proteins play crucial roles in plant growth and development (Kranz et al., 2000; Dubos et al., 2010). With regards to the flower development, *AtMYB33* and *AtMYB65* were microRNA-regulated genes that could promote anther and pollen development (Millar and Gubler, 2005). *AtMYB80* regulated pollen development and programed cell death in the chorioallantoic layer (Phan et al., 2011). *AtMYB106* regulated the flowering time of plants by inhibiting the expression of the *Flowering Locus T* (*AtFT*) gene, which is a negative regulator in plant flowering development (Hong et al., 2021). In chrysanthemum, *CmMYB2* interacted with CmBBX24 to influence the synthesis of gibberellin and ultimately regulated flowering (Zhu et al., 2020). In addition, the R2R3-MYB TFs are known to be widely involved in the regulation of plant specialized metabolism. Previous studies showed that several *AtR2R3-MYB* genes such as *AtMYB11*/*PFG1*, *AtMYB12*/*PFG1*, *AtMYB111*/*PFG3*, *AtMYB21*, and *AtMYB24* were involved in regulating the flavonol biosynthesis pathway in Arabidopsis (Stracke et al., 2007; Dubos et al., 2010; Shan et al., 2020). Overexpression of a flavonol regulator, *MtMYB134*, in hairy roots of *Medicago truncatula* could enhance the biosynthesis of various flavonol derivatives (Naik et al., 2021). *AtMYB75*/*PAP1*, *AtMYB90*/*PAP2*, *AtMYB113*, and *AtMYB114* were identified as regulators of anthocyanin biosynthesis (Teng et al., 2005; Gonzalez et al., 2008). R2R3-MYB family genes are also widely involved in the plant responses to biotic and abiotic stress besides participating in various physiological activities. *AtMYB15* and *AtMYB44* could enhance drought and salt tolerance by regulating stomatal closure under abscisic acid (ABA) treatments (Jung et al., 2008; Ding et al., 2009). *AtMYB96* increased resistance to low temperature through promoting the expression of C-repeat binding factor (*CBF*) (Lee and Seo, 2015). The overexpression of *AtMYB96* also enhanced drought tolerance by mediating ABA signaling and regulating the target gene Lipid transfer protein 3 (*LTP3*) (Seo et al., 2009). In other species, *LcMYB2* promoted seed germination and root growth under drought stress (Zhao et al., 2019). Overexpression of *ZmMYB3R* improved tolerance

to drought and salt stress in transgenic Arabidopsis plants (Wu et al., 2019).

Longan (*Dimocarpus longan* Lour., 2n = 2x = 30) is an important tropical and subtropical fruit tree of the *Sapindaceae* family in the Sapindales order, widely grown in southern China and Southeast Asia (Matsumoto, 2006). As one of the cradles of longan, China currently ranks first in the longan cultivation acreage and yields in the world accounting for 50% of global longan production. Longan has high nutritional and medicinal values and mainly uses its palatable and juicy fruit for commercial purposes (Rangkadilok et al., 2005; Zhang et al., 2020). Flower induction (FI) is an important regulatory process in the transition from vegetative to reproductive growth in plants. Environmental factors like low temperature, drought, and oxidative stress can regulate the flowering of many plants. Potassium chlorate ($KClO_3$) treatment is one of the most effective ways that widely apply in longan production to induce off-season flowering (Huang et al., 2021). To the best of our knowledge, $KClO_3$-induced flowering has only been utilized in longan (Huang et al., 2021). Studies on the physiological mechanism of $KClO_3$-mediated FI showed that $KClO_3$ could increase the content of hormones such as cytokinin (CK) and salicylic acid (SA) in longan, which are closely related to plant flower formation (Sringarm et al., 2009). The $KClO_3$ treatment up-regulated CK concentration together with *DlFT1* expression in mature leaves (Winterhagen et al., 2020). In addition, $KClO_3$ treatment caused changes in the carbon to nitrogen ratio by decreasing the starch content and increasing the soluble sugar, sucrose and soluble amino acid content, which provided nutrients for the flower bud differentiation (Chang, 2010). A regulatory model of FI by $KClO_3$ in longan has been proposed: $KClO_3$ may be reduced to chlorite and hypochlorite under the action of nitrate reductase (NR) and nitrite reductase (NiR) to directly induce stress responses, leading to an increase in CK content and expression of flowering-related genes (Borges et al., 2004; Cho et al., 2017). Thus far, several flowering-related genes have been identified and analyzed in longan such as *flowering locus T* (*DlFT*), *gigantea* (*DlGI*), *early flowering 4* (*DlELF4*), *flavin-binding, kelch repeat, F-box 1* (*DlFKF1*), *short vegetative phase* (*DlSVP*) and *APETALA1-like (DlAP1*) genes (Winterhagen et al., 2013; Jia et al., 2014; Huang F. N. et al., 2017; Waheed et al., 2020) but the molecular mechanism underlying longan FI and floral organ development is largely unknown.

The *R2R3-MYB* genes comprise one of the largest TF families in longan. Previous studies have identified 35 (Zheng et al., 2020) and 98 (Lin et al., 2017) *R2R3-MYB* genes based on transcriptome data and the Next Generation Sequencing (NGS)-based assembled genome sequences, respectively. Recently, a high-quality chromosome-level genome of longan "Shixia" ("SX") was assembled by our team (NCBI GenBank: JAIFKA000000000.1), which will help uncover the distribution, evolution and function of genome-wide R2R3-MYB TFs. By combining PacBio long reads, HiC data and genetic maps, the longan genome was assembled into 483.32 Mb with a contig N50 size of 763.91 kb, which were further anchored onto 15 pseudochromosomes with a 99.30% anchored rate. Benchmarking Universal Single-Copy Orthologs (BUSCO) assessment of the genome was 94.6%, which was an indicator of good quality assembly. Transcriptome dynamics revealed that six associated genes including one *MYB* transcript factor (*Dil.04g005810.2.t1*) were closely related to the flowering time under the $KClO_3$ treatment (unpublished data). Studying the *MYB* gene expression profiles associated with off-season FI by $KClO_3$ in longan will provide a deeper understanding of the molecular regulation mechanism of the flowering process in woody plants.

In this present study, we identified 119 *DlMYB* genes from the "Shixia" longan genome sequences and divided them into 28 subgroups. The *DlMYB* gene sequences were further investigated for their characterization, phylogenetic relationship, gene structure, subcellular localization, conserved motifs, *cis*-acting elements, chromosome distribution, synteny analysis, selection pressure, and gene duplication. In addition, the RNA-seq data showed expression patterns of *DlMYBs* in nine different tissues and in flower bud tissues treated with $KClO_3$ at different stages which assisted us in determining the potential role of *DlMYBs* in longan flowering. The qRT-PCR method was used to verify the expression levels of the selected *MYB* genes in longan tissues. Moreover, three potential flowering-associated *MYB* genes were further experimentally validated by their subcellular localization in Arabidopsis protoplasts. This is the first systematic report on the genome-wide characterization of the *MYB* gene family in longan. This study will help investigate members of the R2R3-MYB family involved in the flowering process in longan and the regulatory network, which will pave the way for future functional studies and candidate gene selection for genetic improvement in longan.

MATERIALS AND METHODS

Plant Materials and $KClO_3$ Treatments

Longan cultivar trees ("Shixia," abbreviated "SX") about 4— to 6— year-olds in Maoming, Guangdong, China, were selected for treatment with white $KClO_3$ powder of >99% purity in 2016. The apical buds were collected under controlled conditions and $KClO_3$ treatment from longan trees at ten time points with three biological replications including 0 day after treatment (DAT) (November 18th, 2016), 5 DAT (November 23th, 2016), 10 DAT (November 28th, 2016), 15 DAT (December 3rd, 2016), 20 DAT (December 8th, 2016), 25 DAT (December 13th, 2016), 30 DAT (December 18th, 2016), 35 DAT (December 23th, 2016), 41 DAT (December 29th, 2016) and 54 DAT (January 11th, 2017). Immediately after harvesting, all tissue samples were frozen in liquid nitrogen and then stored at −80°C prior to the RNA isolation.

Illumina Library Construction and RNA Sequencing

Total RNAs were extracted from collected samples with three biological replicates using the RNAprep Pure Plant Kit (DP432, TIANGEN Biotech, Beijing, China) according to the manufacturer's protocol. The quality of total RNAs was verified on an Agilent 2100 Bioanalyzer (Plant RNA Nano

Chip, Agilent, CA, United States). After quantity and quality determination, RNA-Seq libraries were constructed using the Illumina® TruSeq™ RNA Sample Preparation Kit [RS-122-2001(2), Illumina] and then subjected to Illumina transcriptome short reads sequencing on Illumina X Ten platform with paired-end (PE) 150 nt mode (San Diego, CA, United States).

Public Data Retrieval

The genomic data of *Dimocarpus longan* "Shixia" (GenBank accession No.: JAIFKA000000000.1), *Litchi chinensis* (GenBank accession No.: JAIUGE000000000.1), *Oryza sativa* (GenBank accession No.: RPSM01000000), and *Zea mays* (GenBank accession No.: CABHLF010000000) were downloaded from the National Center for Biotechnology Information (NCBI). The data for *Xanthoceras sorbifolium* was downloaded from GigaDB (Sneddon et al., 2012). The data for *Arabidopsis thaliana* R2R3-MYB genes were retrieved from The Arabidopsis Information Resource (TAIR) database. The RNA-seq data of "Sijimi" ("SJM") longan were downloaded from the NCBI (GenBank accession No.: PRJNA329283), including nine tissues (roots, stems, leaves, flower buds, flowers, young fruits, pulps, pericarps, and seeds) from a previous study (Lin et al., 2017).

Identification of Longan R2R3-MYB Family Genes

The Hidden Markov Model (HMM) profiles for the Myb_DNA-binding domain (PF00249) and SWIRM domain (PF04433) were downloaded from the Pfam protein family database. The specific identification method of *R2R3-MYB* genes was as follows. The largest number of R2R3-MYB proteins were identified based on the HMM profiles of the MYB domain and SWIRM domain by using HMMER 3.0 with *E*-values $\leq$ 0.01 threshold (Potter et al., 2018). Gene sequences containing MYB domains but no SWIRM domain were considered as putative *MYB* genes. All longan *R2R3-MYB* genes were verified using *AtR2R3-MYB* gene sequences as target sequences by BLASTP, with the following parameters: *E*-value cut-off of $1e^{-5}$ and 30% identity. These sequences were manually inspected by the NCBI Conserved Domain Database (NCBI-CDD) and the SMART website to ensure that the genes contained two MYB domains. The full-length amino acid sequences of DlMYB proteins were submitted to the Expasy website (Gasteiger et al., 2003) to predict their molecular weights (MW), isoelectric points (PIs), hydrophilia, and instability. Detailed information on the *DlMYB* genes is listed in **Supplementary Table 1**. Multiple sequence alignments of the MYB domains were performed by Clustal X (2.1) (Larkin et al., 2007) with default parameters. The amino acid sequences in MYB domains were visualized by GeneDoc (Nicholas et al., 1997). Weblogo (Crooks et al., 2004) was used to display the sequence logos of R2 and R3 MYB domain repeats based on the multiple alignment files.

Phylogenetic Analysis

The multiple sequence alignments of the full-length R2R3-MYB protein sequences from longan and Arabidopsis were performed by Muscle v3.8.1 (Edgar, 2004) with default parameters. The phylogenetic tree of R2R3-MYB proteins from both genomes was constructed using the Maximum Likelihood (ML) method of RAxML (Stamatakis, 2014), with the following parameters: JTT model and 1,000 bootstrap replications. The Evolview (Subramanian et al., 2019) was used to visualize the ML phylogenetic tree. Longan R2R3-MYB proteins were grouped according to the previously reported *AtR2R3-MYB* family classification (S1–S25) (Stracke et al., 2001; Dubos et al., 2010). Another ML phylogenetic tree of R2R3-MYB protein sequences from the longan genome was constructed using the same method.

Motif Prediction, Gene Structure, and *Cis*-Acting Element Analysis

The MEME v5.3.3 (Multiple expectation maximization for Motif Elicitation) (Bailey et al., 2009) online tool was used to determine the conserved motifs in DlMYB amino acid sequences, with the following parameters: zero or one occurrence per sequence; maximum number of motif to find, 20; minimum width of motif, 6; and maximum width of motif, 200. Weblogo (Crooks et al., 2004) was used to display these conserved motifs.

The exon/intron structure patterns of *DlMYB* genes, including intron distribution patterns and intron-exon boundaries, were graphically visualized by the TBtools v1.087 (Chen et al., 2020), using the nucleotide sequences of *DlMYB* genes and CDSs (coding sequences). The 2-kb sequences upstream of each longan *R2R3-MYB* gene sequence were submitted to the PlantCARE database (Lescot et al., 2002) to predict *cis*-acting elements. The predicted results of gene structures and *cis*-acting elements were visualized by TBtools program (Chen et al., 2020).

Chromosomal Locations and Synteny Analysis

The detailed information of the physical locations of *DlMYB* genes were generated based on the longan genome sequences (NCBI GenBank: JAIFKA000000000.1). The TBtools was used to exhibit chromosomal locations of *DlMYB* genes using the gene location visualization program. The Multiple Collinearity Scan toolkit (MCScanX) (Wang et al., 2012) and TBtools were used to analyze and visualize *DlMYB* gene replication events with *E*-value cut-off of $< 1e^{-10}$. The Dual Systeny Plot program in TBtools (Chen et al., 2020) was adopted to elucidate the synteny relationship of orthologous *R2R3-MYB* genes between longan and other five species, including three dicots Arabidopsis, litchi and yellow horn, and two monocots rice and maize. The Ka/Ks (non-synonymous/synonymous substitution ratio) values were used to determine the selection pressure after replication. Ka/Ks $>$ 1 signifies purified selection, Ka/Ks = 1 indicates neutral selection, and Ka/Ks $<$ 1 represents positive selection (Hurst, 2002). The MS model in KaKs_calculator 2.0 (Wang et al., 2010) was used to calculate the values of Ka, Ks, and Ka/Ks of *MYB* duplicated gene pairs.

Subcellular Localization Analysis

Subcellular localization was used to analyze the specific location of the *DlMYB* genes within the cell. The subcellular localization and nuclear localization signal (NLS) were predicted using

the YLoc (Briesemeister et al., 2010) and NLStradamus online tool (Nguyen Ba et al., 2009), respectively. Three flower-preferential expressed *DlMYB* genes, *DlMYB16*, *DlMYB72*, and *DlMYB116* were selected for further experimental validation of their subcellular localizations. Total RNAs were extracted from leaves, flowers and flower buds of "Shixia" longan plants using the RNAprep pure Plant Kit (Tiangen, #DP432) according to the manufacturer's protocol. The full-length coding sequences excluding stop codon of these three *R2R3-MYB* genes were amplified and cloned into the *Bsa*I/*Eco*31I restriction sites of the pBWA(V)HS-GLosgfp vector (**Supplementary Figure 5**; Sun et al., 2018). Briefly, the pBWA(V)HS-DlMYB-GFP (35S:GFP) recombinant vector and the nucleus marker vector pBWA(V)HS-Nucleus-mKATE (Shcherbo et al., 2007) were co-transformed in Arabidopsis protoplasts with PEG solution (40% W/V PEG 4000, 0.6 mol/L Mannitol, 0.1 mol/L calcium chloride, pH 5.8). The pBWA(V)HS-GFP empty vector was used as a control. After incubation in darkness at 28°C for 18–24 h, the subcellular distribution and localization of DlMYB-GFP were captured in protoplasts using the Nikon C2-ER laser scanning confocal microscope (LSCM) at a 488 nm excitation wavelength and a 510 nm emission wavelength. Meanwhile, colocalization of DlMYB-GFP and Nucleus-mKATE (NLS-mKate) was used to confirm the nuclear localization of MYB proteins with the mKATE protein at 561 nm excitation wavelength and 580 nm emission wavelength (Shcherbo et al., 2007; Zhao et al., 2017). The chlorophyll fluorescence was captured to confirm the localization of chloroplasts at 640 nm excitation wavelength and 675 nm emission wavelength. The primers used for PCR amplification are listed in **Supplementary Table 2**.

Gene Expression Pattern Analysis by RNA-seq

RNA-seq data from nine tissues of "Sijimi" ("SJM") longan (See more details in section "Public Data Retrieval") were used to explore the expression patterns of *R2R3-MYB* genes at different developmental stages of the longan plant (Lin et al., 2017). The Fastp v0.20.1 software (Chen et al., 2018) was used to filter low-quality reads and adaptor sequences. High-quality clean reads were mapped to the "SX" longan reference genome sequences using HISAT2 v2.2.1 (Kim et al., 2015). The trimmed mean of *M* values (TMM) method was used to normalize the expression levels of *R2R3-MYB* genes in longan using edgR software (Robinson et al., 2010). The expression profiles of the *R2R3-MYB* genes in longan tissues were visualized by TBtools and the ratios were $\log_2$-transformed. Gene ontology (GO) annotations were conducted in the eggNOG-mapper (Huerta-Cepas et al., 2019) for the *DlMYB* genes. All 37,142 genes of longan were used as the reference set, and 119 *DlMYB* genes were taken as the test set. The annotation results of *DlMYB* genes were enriched and analyzed using the R package clusterProfiler (Yu et al., 2012), including three categories, cellular component (CC), biological process (BP), and molecular function (MF).

The RNA-seq data of "SX" under $KClO_3$ treatment was also processed using the same method. The differentially expressed genes (DEGs) under $KClO_3$ treatments and controlled conditions in longan were identified using the R package DESeq2 v3.11 (Love et al., 2014), with the threshold of $|\log2(\text{fold-change})| \geq 1$ and a false discovery rate (FDR) ≤ 0.05. The DESeq2's Likelihood Ratio Test was used to conduct the time course analysis to find genes which were significantly upregulated/downregulated over time in different treatments. The weighted gene co-expression network analysis (WGCNA) was conducted using the WGCNA v4.0.2 package (Langfelder and Horvath, 2008) based on DEGs with TMM ≥ 1 at least in one sample. Cytoscape v3.7.1 software (Shannon et al., 2003) was used to visualize the gene interaction networks of longan *R2R3-MYB* genes in modules. FLOweRing Interactive Database (FLOR-ID) of Arabidopsis (Bouché et al., 2016) was used to search flowering-related genes that may interact with *DlR2R3-MYB* genes. GO and KEGG enrichment analysis were performed using the R package clusterProfiler.

Expression Validation by Quantitative Real-Time PCR

Eight *DlMYB* genes potentially involved in the flower development were selected for the quantitative real-time polymerase chain reaction (qRT-PCR) in five different tissues (leaves, pericarps, pulps, flowers, and flower buds) of "Shixia." The total RNA in the samples were isolated using the TRNzol Universal (DP424, TIANGEN Biotech, Beijing, China) reagent according to the manufacturer's instructions. The concentration and quality of the extracted RNAs was determined using Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, MA, United States) and gel electrophoresis. For cDNA synthesis, 0.5 μg RNA was reverse-transcribed using the M-MLV 4 First-Strand cDNA Synthesis Kit (MT403-01, Biomed, Beijing, China) following the supplier's instructions. The cDNA was stored at −20°C for further use. Gene-specific primers were designed using the online tool primer3plus (Untergasser et al., 2007; **Supplementary Table 2**). The longan actin gene (*Dil.06g016430.1*) was used as the internal reference gene. The qRT-PCR amplification was conducted on a CFX Connect Real-Time PCR detection system (Bio-Rad, Hercules, CA, United States) with 2 × SYBR Green qPCR MasterMix (MT521, Biomed, Beijing, China). Three independent biological replicates and three technical repeats were taken. The reaction was performed using the protocol for this kit with the following minor modifications: 95°C for 30 s, 40 cycles of 95°C for 5 s, and 60°C for 30 s. The relative expression levels of *DlMYB* genes were further calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

RESULTS

Identification and Characterization of Longan *MYB* Genes

A total of 219 non-redundant *DlMYB* genes were identified in the longan reference genome, including 119 *R2R3-MYB* genes (2R-MYBs), 95 *MYB-like* repeats genes (1R-MYBs), three *R1R2R3-MYB* genes (3R-MYBs), and two *4RMYB* genes (4R-MYBs). The *R2R3-MYB* genes constituted the largest group of the MYB gene family in longan. The *R2R3-MYB* genes were renamed *DlMYB1* to *DlMYB119* based on their chromosomal locations in longan

genome (**Supplementary Table 1**). The detailed information of *DlMYB* gene features, including coding sequences, protein sequences, chromosomal locations, protein sizes, molecular weights (MWs), isoelectric points (PIs), subcellular localizations, instability, and hydrophilia, were listed in **Supplementary Tables 1, 3**. The length of DlMYB proteins ranged from 134 (DlMYB38) to 661 (DlMYB55) amino acids (aa) with an average of 306 aa, and the MWs had a range of sizes from 15.7 (DlMYB55) to 75.8 (DlMYB38) KDa (mean 34.6). The PIs of the DlMYB proteins ranged from 4.91 (DlMYB32) to 10.60 (DlMYB23). All DlMYB protein hydrophilia values were negative, indicating that longan MYBs are hydrophilic.

The 119 R2R3-MYB amino acid sequence logos of R2 and R3 repeats were produced to show the sequence characterization and frequency of the most widespread amino acids at each position (**Figure 1** and **Supplementary Figure 1**). In general, the R2 and R3 repeats of MYB domains had ∼108 basic residues (including the linker region). Critical amino acid insertions were detected in both R repeats, especially in the Helix2 (H2) segment of R2 and R3. The R2R3-MYB domains showed amino acid insertion points for DlMYB23 (24aa), DlMYB27 (8aa), DlMYB98 (3aa), and DlMYB101 (3aa) in R2, DlMYB4 (2aa), DlMYB53 (2aa), and DlMYB54 (2aa) in R3 (**Figure 1** and **Supplementary Figure 1**). The results confirmed the conservation of the MYB domain, but amino acid compositions and lengths of the regions outside the MYB domain were highly variable.

A series of highly conserved and regularly distributed Trp (W) residues were detected in R2 and R3 repeats of DlMYBs, which are noted for landmarks of plant MYB proteins. Interestingly, only the second and third Trp residues were found to be conserved in the R3 repeat and the first Trp was generally replaced by a hydrophobic amino acid, such as Phe (F) or Ile (I). Apart from Phe and Ile, the substitution with the amino acids Met (M), Leu (L), and Try (Y) were also found in the first Trp residue in the R3 repeat at Trp-62 position of DlMYB proteins (DlMYB6, DlMYB10, DlMYB51, DlMYB57, DlMYB72, DlMYB81, DlMYB82, DlMYB98, DlMYB101, and DlMYB114). In addition, the Trp residues, Cys-45, Arg-48 in the R2 repeat, Leu-53 in the linker region, and Glu-66, Gly-78 in the R3 repeat were conserved in all DlMYB proteins (**Figure 1** and **Supplementary Figure 1**). These highly conserved amino acid residues were mainly distributed in the third helix and the helix-turn-helix (HTH) motif. In DlMYB32 and DlMYB75, the Pro-55 was replaced with Ala and Ser in the R3 repeat, respectively, implying they may have different evolutionary divergence rates.

Phylogenetic Analysis and Classification of Longan R2R3-MYB Proteins

The maximum likelihood (ML) phylogenetic tree of 119 DlMYB and 126 AtMYB proteins showed that all R2R3-MYBs were categorized into 28 subgroups (designed as S1–S28) based on sequence similarity and topology (**Figure 2**). Each clade in the ML tree contained R2R3-MYB proteins from longan and Arabidopsis. Five AtMYBs and thirteen DlMYBs that could not be retrieved in the previously constructed AtMYB phylogenetic tree were rearranged in this study, in S26–S28, according to the phylogenetic analysis of Arabidopsis and longan MYBs. Eventually, all 119 longan R2R3-MYB proteins were divided into 28 subgroups, including S1 (three members), S2 (seven members), S3 (one member), S4 (seven members), S5 (nine members), S6 (eight members), S7 (one member), S8 (four members), S9 (four members), S10 (two members), S11 (three members), S12 (one member), S13 (seven members), S14 (seven members), S15 (four members), S16 (three members), S17 (four members), S18 (four members), S19 (one member), S20 (four members), S21 (five members), S22 (five members), S23 (one member), S24 (five members), S25 (six members), S26 (four members), S27 (four members), and S28 (five members) (**Figure 2** and **Supplementary Table 1**). Not only the gene duplications but gene loss events seem to have an impact on the expansion and contraction of MYB members after the divergence of longan and Arabidopsis genomes. As shown in **Figure 2**, some subgroups (S2, S5, S6, S24, S27, S28) of the phylogenetic tree harbored twice as many longan MYB members as Arabidopsis MYBs. The opposite was also observed in subgroups 7, 10, 12, 16, 19, and 23, where the number of DlMYB proteins was less than half that of AtMYB proteins.

Motif Composition, Gene Structure, and Promoter *Cis*-Acting Element Analysis

The ML phylogenetic tree constructed for 119 DlMYB proteinsalone also revealed that all DlMYB proteins were divided into 28 subgroups (**Figure 3A**). In order to have a more comprehensive understanding of the diversity and conserved domains of these *R2R3-MYB* genes in longan, conserved motifs of these DlMYB protein sequences were examined. A total of 20 conserved motifs were examined and designated as motif 1 to 20 (**Figure 3B**, **Supplementary Figure 2**, and **Supplementary Table 4**). The majority of R2R3-MYB protein sequences had four significantly similar conserved motifs and motif orders. Ninety-seven out of 119 R2R3-MYB protein sequences contained motif 3, motif 4, motif 1, and motif 2 present in that order, and these motifs were composed of 21, 11, 40, and 29 amino acids, respectively (**Figure 3B** and **Supplementary Table 4**). Those four motifs were found to encode the MYB DNA-binding domain. Motifs 3, 4, and the left part of motif 1 were composed of the R2 repeat, and the right part of motif 1 and motif 2 were found in the R3 repeat. In other 17 R2R3-MYB proteins, motif 4 in the R2 repeat was replaced by motif 18. In addition, motif 4 and motif 1 of DlMYB28 and DlMYB29 were replaced by motif 8. Three R2R3-MYB proteins, DlMYB98, DlMYB100, and DlMYB101 only contained motif 1, motif 2, and motif 3. Almost all DlMYB proteins (118/119) contained motif 2 and motif 3 with the exception of DlMYB 42 that did not have motif 2. Most DlMYB proteins from the same clade had a similar motif composition. For instance, motifs 9 and 17 were shared by the members in S6, motifs 8 and 13 were exclusively in S27, and motif 15 was unique to S25, suggesting that these unique motifs may contribute to the differentiation in specific functions.

The exon/intron organization was analyzed to investigate the structural diversity of *DlR2R3-MYB* genes. The number of introns in longan *R2R3-MYB* genes ranged from zero to

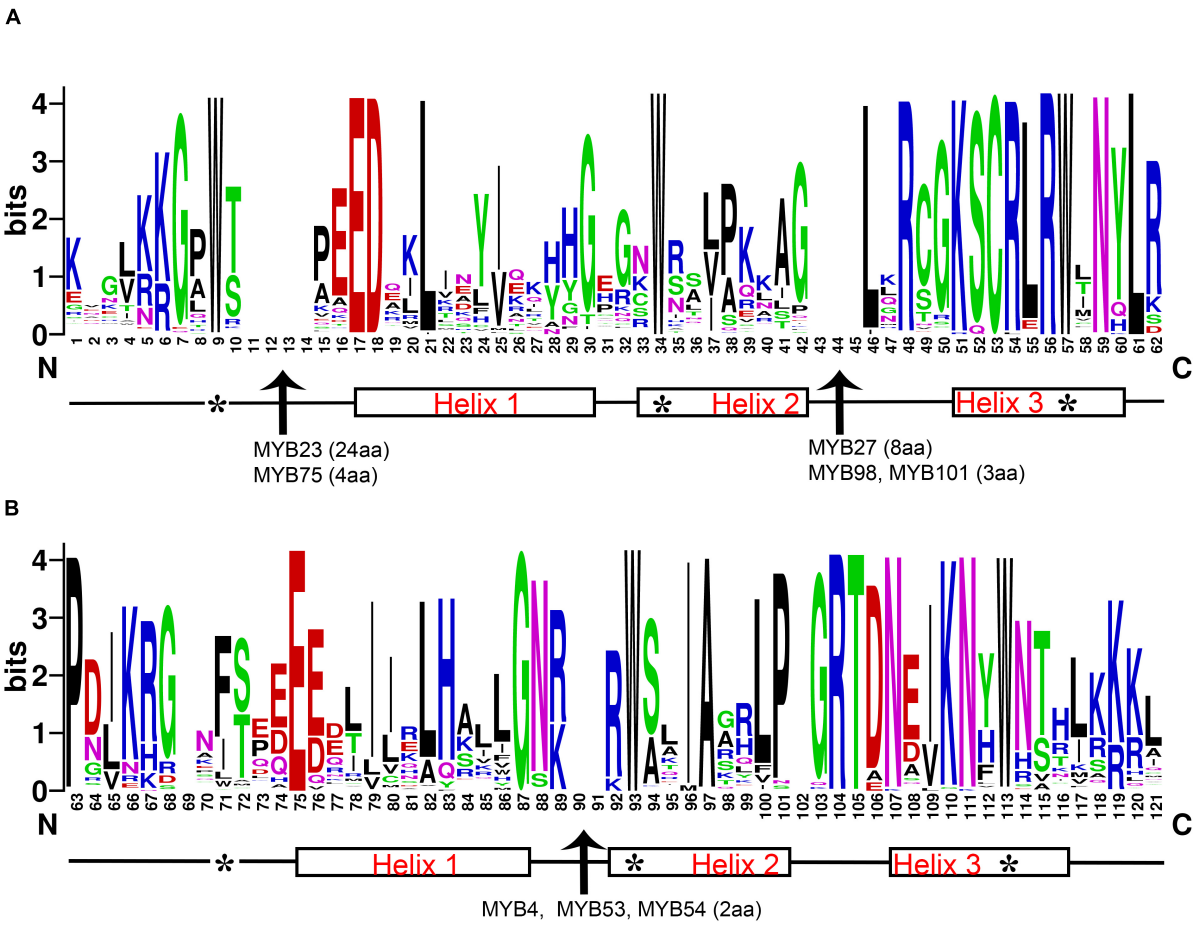


FIGURE 1 | Characterization of R2 **(A)** and R3 **(B)** MYB repeats of longan R2R3-MYB proteins. The position score indicated the degree of conservation of each amino acid at the specific position within sequences. The positions of the three α-helices are marked (helices 1–3). Arrows represent amino acid insertion sites detected in the R2 repeat of genes *DlMYB23*, *DlMYB27*, *DlMYB75*, and *DlMYB98* and the R3 repeat of *DlMYB53*, *DlMYB54*, and *DlMYB93*. The length of each insertion (amino acid, aa) is highlighted in parentheses. Characteristic tryptophan (W) residues are marked with the symbol "*".

thirteen (**Figure 3C**). Most (111/119) of the coding sequences of *R2R3-MYB* genes in longan were disrupted by one to three introns, with the exception of six *DlMYB* genes (*DlMYB36*, *DlMYB43*, *DlMYB67*, *DlMYB68*, *DlMYB78*, and *DlMYB101*) that completely lacked introns, indicating the conserved intron number in the *R2R3-MYB* genes. Only two *DlMYB* genes had an excess of introns, including *DlMYB75* (13 introns) and *DlMYB55* (8 introns). Most genes clustered in the same subgroup exhibited highly similar gene structures and intron numbers, which could strongly support the accuracy of the classification of subgroups.

The predicted results for *cis*-regulatory elements showed that a total of fourteen types of *cis*-regulatory elements were predicted in *DlR2R3-MYB* genes. The most widespread *cis*-acting elements in the promoter region of the *DlMYB* genes were light responsiveness (119/119), anaerobic induction (102/119), and ABA responsiveness elements (101/119) (**Figure 3D** and **Supplementary Tables 5, 6**), while the least common ones were wound-responsive (4/119), flavonoid biosynthetic genes regulation (13/119) and circadian (23/119). Moreover, the promoters of 79 (66.39%), 70 (58.82%), and 56 (47.06%) *DlMYB* genes were predicted to contain *cis*-acting elements associated with MeJA (methyl jasmonate), GA (gibberellin), and SA (salicylic acid), respectively. The detailed information regarding the *cis*-regulatory elements in longan *DlMYB* genes is given in **Supplementary Tables 5, 6**. All 14 types of *cis*-acting elements could be divided into three major functional categories, including cellular development or photoresponsive elements (two types of elements), phytohormone (five types of elements), and environmental stress (seven types of elements). In general, *DlR2R3-MYB* gene promoters containing environmental stress-related elements (398) were the most common followed by phytohormone-related ones (358).

Chromosomal Distribution, Duplication Events and Interspecies Synteny Analysis

Genome chromosomal location analysis showed that 119 *R2R3-MYB* genes were unevenly distributed on 15 chromosomes (**Figure 4A**), and each chromosome had at least two

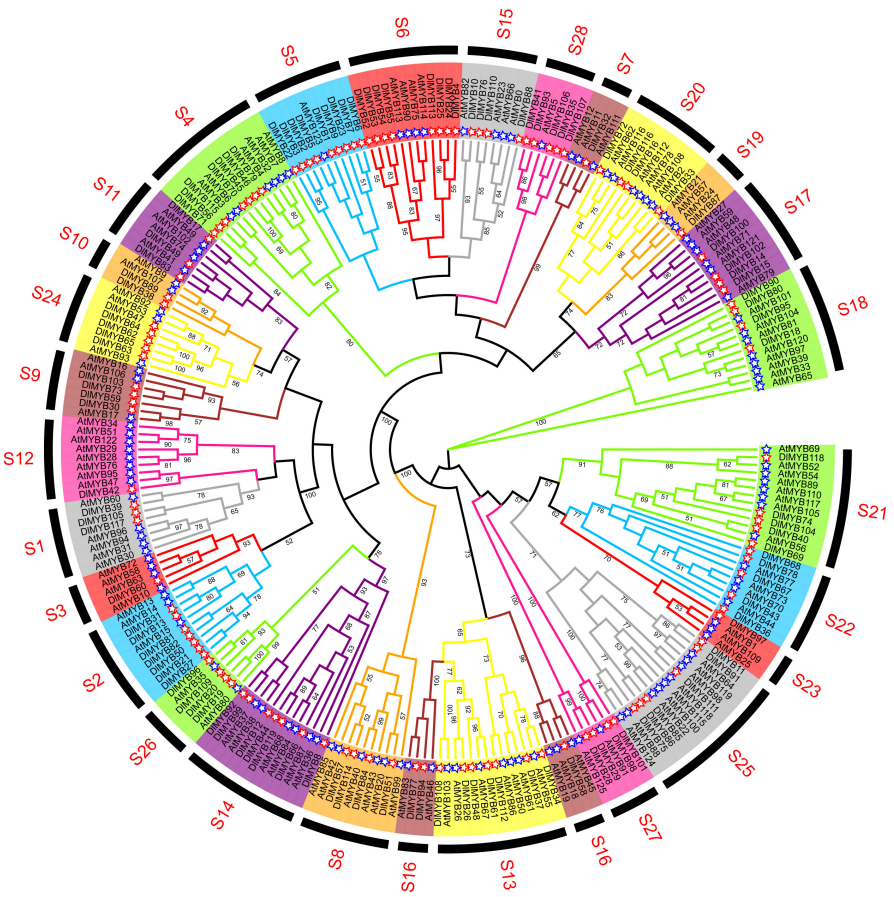


FIGURE 2 | Phylogenetic analysis of longan and *Arabidopsis* MYB proteins. The full-length amino acid sequences of longan (119) and Arabidopsis (126) MYB proteins were compared using RAxML software to generate ML phylogenetic trees with 1,000 bootstrap values. The Evolview3 was used to visualize and display > 50 bootstrap values at the nodes. Red and blue stars represent DlMYB and AtMYB proteins, respectively. All MYB proteins were divided into 28 subgroups (S1–S28).

DlR2R3-MYB genes. DlChr1 harbored the largest group of *R2R3-MYB* genes (16/119), followed by DlChr15 with 14 *R2R3-MYB* genes and DlChr12 with 11 genes, while only two genes were located on DlChr11. *DlR2R3-MYB* genes tend to cluster to the proximal end of some chromosomes, including DlChr1, DlChr2, DlChr3, DlChr4, DlChr8, DlChr10, and DlChr12.

In our study, a total of 25 segmental duplications (SD) involving 44 *MYB* genes were identified on different chromosomes (**Figure 4B** and **Supplementary Tables 7, 8**). *DlMYB* genes were located within synteny blocks on almost all chromosomes except for DlChr5. The SD events occurred mostly on chromosomes 1, 12, and 15 (**Supplementary Table 8**). Some *DlR2R3-MYB* genes, such as *DlMYB2*, *DlMYB78*, and *DlMYB110*, were involved in multiple duplications. In total, six tandem duplications (TD) clusters consisting of 16 *DlR2R3-MYB* genes were identified in DlChr1, DlChr8, DlChr12, and DlChr15 (**Figure 4A**). The distribution of TD genes on chromosomes were consistent with aforementioned chromosomes with a high-density of *MYB* genes. These results indicated that approximately 43.7% of *DlMYB* genes might be generated by duplication events, which played a vital role in the expansion of the *MYB* gene family in longan (**Supplementary Table 7**). The *Ka*/*Ks* values (non-synonymous/synonymous substitution ratios) were calculated to understand the selection mode of segmentally and tandemly duplicated *R2R3-MYB* gene pairs in longan. *Ka* had a range of values from 0.06 to 0.52 while *Ks* values varied from 0.14 to 4.36 (**Supplementary Table 8**). In particular, one duplicated gene pair *DlMYB14-DlMYB15* had no *Ka* and *Ks* value (NaN). Through manual checking, we found that mutations between *DlMYB14* and *DlMYB15* sequences occurred at the nucleic acid level but their amino acid sequences remained unchanged. The *Ka*/*Ks* ratios of 35 *DlMYB* gene duplicated pairs (including 10 tandem and 25 segmental duplication gene pairs) ranged from 0.05 to 0.84 (**Supplementary Table 8**) (all less than 1), indicating that longan *R2R3-MYB* duplicated genes had undergone negative purifying selection in the process of evolution and contribute largely to the maintenance of function in the longan *R2R3-MYB* gene family. The average *Ka*/*Ks* value of tandem duplication genes was 0.34, two-fold that of segmented duplication genes (0.17) (**Supplementary Figure 6**). Furthermore, the *Ka*/*Ks* value of each orthologous *DlMYB* gene pair was also calculated in the *Sapindaceae*

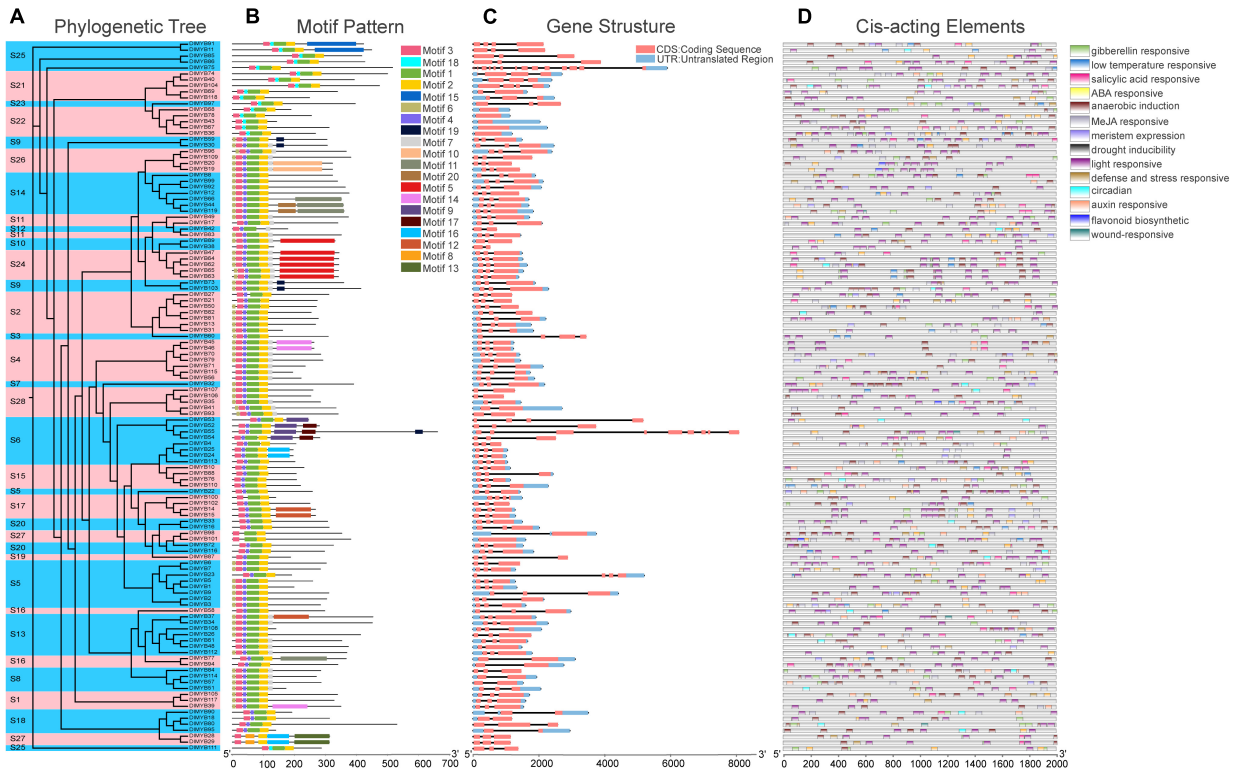


FIGURE 3 | Phylogenetic tree, conserved motifs, gene structural, and *cis*-acting element analysis of DlMYBs. **(A)** An Maximum Likelihood (ML) tree of 119 DlMYBs proteins with 1,000 bootstrap replicates. **(B)** The conserved motifs in DlMYBs were predicted using the MEME Suite web server. **(C)** Exon/intron structures of the longan *R2R3-MYB* genes. The introns, exons, and UTRs are indicated by orange boxes, black lines, and blue boxes, respectively. **(D)** The *cis*-acting elements of the longan R2R3MYB proteins were predicted using PlantCARE.

family (**Supplementary Figure 6** and **Supplementary Table 9**). Nearly all of the 289 (289/291) orthologous gene pairs between longan and litchi, longan and yellow horn had Ka/Ks values of < 1, except for two gene pairs *DlMYB32-LITCHI026531.m1* (Ka/Ks = 1.20) and *DlMYB100-LITCHI018670.m1* (Ka/Ks = 1.78) between longan and litchi, implying that purifying selection may be the dominant force driving the evolution of *R2R3-MYB* genes in the lineage of *Sapindaceae*. The Ka/Ks mean values of the *R2R3-MYB* syntenic gene pairs between longan and Arabidopsis, longan and litchi, longan and yellow horn orthologs were 0.15, 0.32, and 0.23, respectively (**Supplementary Figure 6** and **Supplementary Table 9**).

To gain more insight into the evolutionary model of the longan *R2R3-MYB* gene family, the comparative syntenic maps of longan along with three representative dicots including Arabidopsis, litchi (*Litchi chinensis*) and yellow horn (*Xanthoceras sorbifolium* Bunge), and two monocots, rice (*Oryza sativa* L.) and maize (*Zea mays*) were generated (**Figure 5A**). Litchi and yellow horn were phylogenetically closer to longan, which were also in the family of *Sapindaceae*. The results revealed that 99 *DlMYB* genes were involved in the formation of a syntenic relationship with those in litchi, followed by yellow horn (94 genes involved), Arabidopsis (61), maize (17), and rice (16) (**Figure 5B**). The maximum number of ortholog *DlR2R3-MYB* gene pairs were found in *DlMYB-XsMYB* (148) and *DlMYB-LcMYB* (143) followed by *DlMYB-AtMYB* (102), while the minimum number were found in *DlMYB-ZmMYB* (25) and *DlMYB-OsMYB* (17) (**Supplementary Table 9**), which is consistent with their taxonomic positions. Notably, a total of nine *DlMYB* genes were predicted to form collinear pairs with genes from all the other five species (**Supplementary Table 10**). Eighty-five syntenic gene pairs were only found between longan and other dicots (**Figure 5B**). For example, the gene pair *DlMYB38-AT3G02940.1* was only detected between longan and Arabidopsis, *DlMYB51-EVM0018212.1-AT1G66230.1* was only found among longan, Arabidopsis, and yellow horn. A total of 26 MYB genes were only found in longan, litchi, and yellow horn (**Supplementary Table 10**). Additionally, 18 *AtMYB* genes were predicted to be paired with at least two longan genes (**Supplementary Table 9**), suggesting the newly evolved functions of redundant *DlMYB* genes. The detailed information of syntenic gene pairs is provided in **Supplementary Tables 9, 10**.

Subcellular Localization of DlMYB Proteins

Subcellular localization predictions revealed that all 119 DlMYB proteins were located in the nucleus (**Supplementary Figure 7** and **Supplementary Tables 1, 3**). To experimentally validate the predicted subcellular localization, we selected three potential

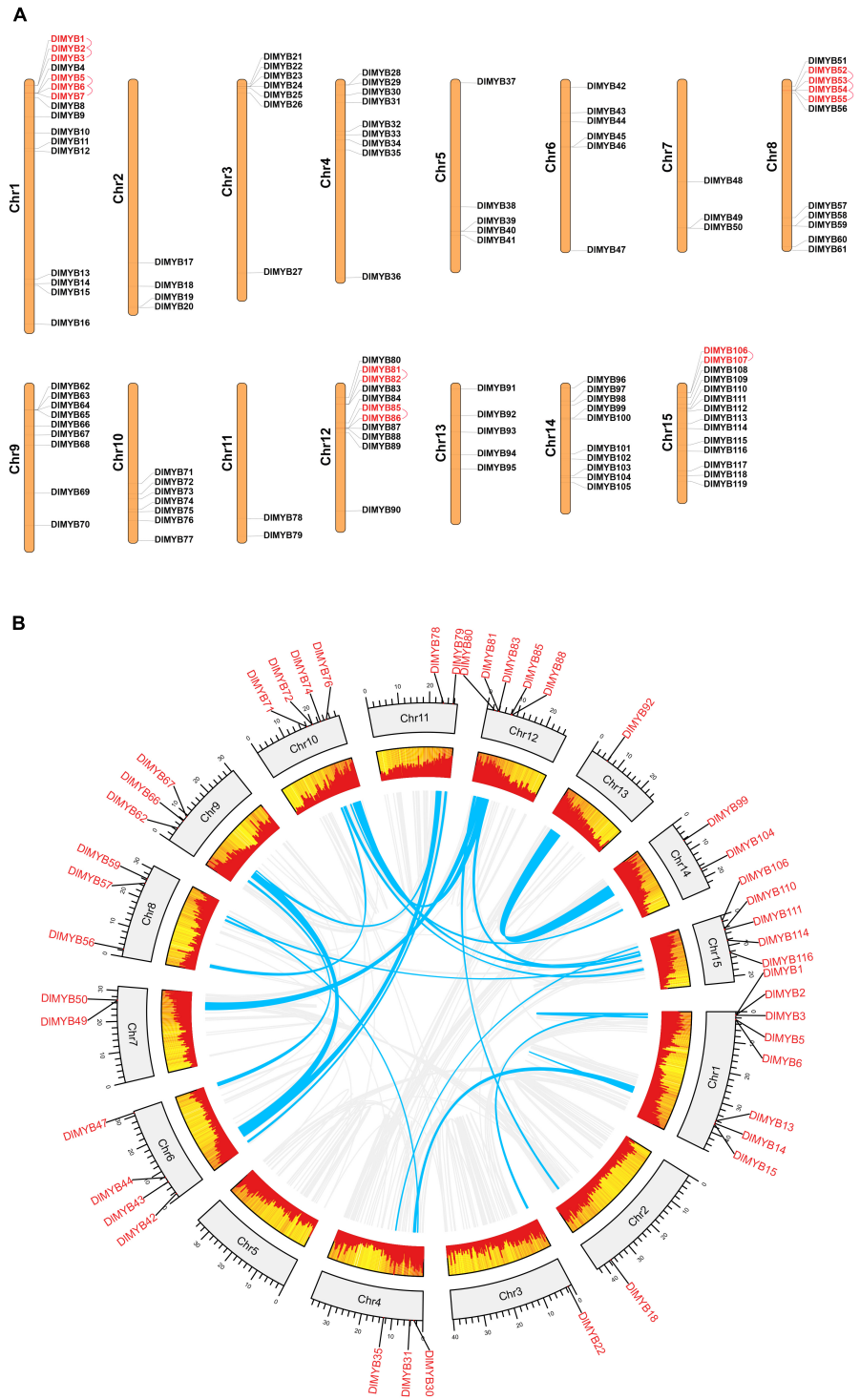


FIGURE 4 | Chromosomal location, tandem duplication, and segmental gene duplication analysis of the *DlMYB* genes. **(A)** The distribution map of *DlMYB* genes on the 15 chromosomes of longan (genes with tandem clusters are marked in red). **(B)** Circos plot of *DlMYB* genes in the longan genome. Gene pairs with segmental duplications (SDs) are indicated by blue lines, and gene density was highlighted by red rectangular boxes.

flowering-associated *DlMYB* genes (*DlMYB16*, *DlMYB72*, and *DlMYB116*) for analysis in Arabidopsis protoplasts. These three genes were expressed at significantly high levels in flower or flower bud tissues compared with other tissues. The NLS-mKATE and Cy5 were applied in each transformed design as a marker for nuclear and chloroplast localization, respectively. The results

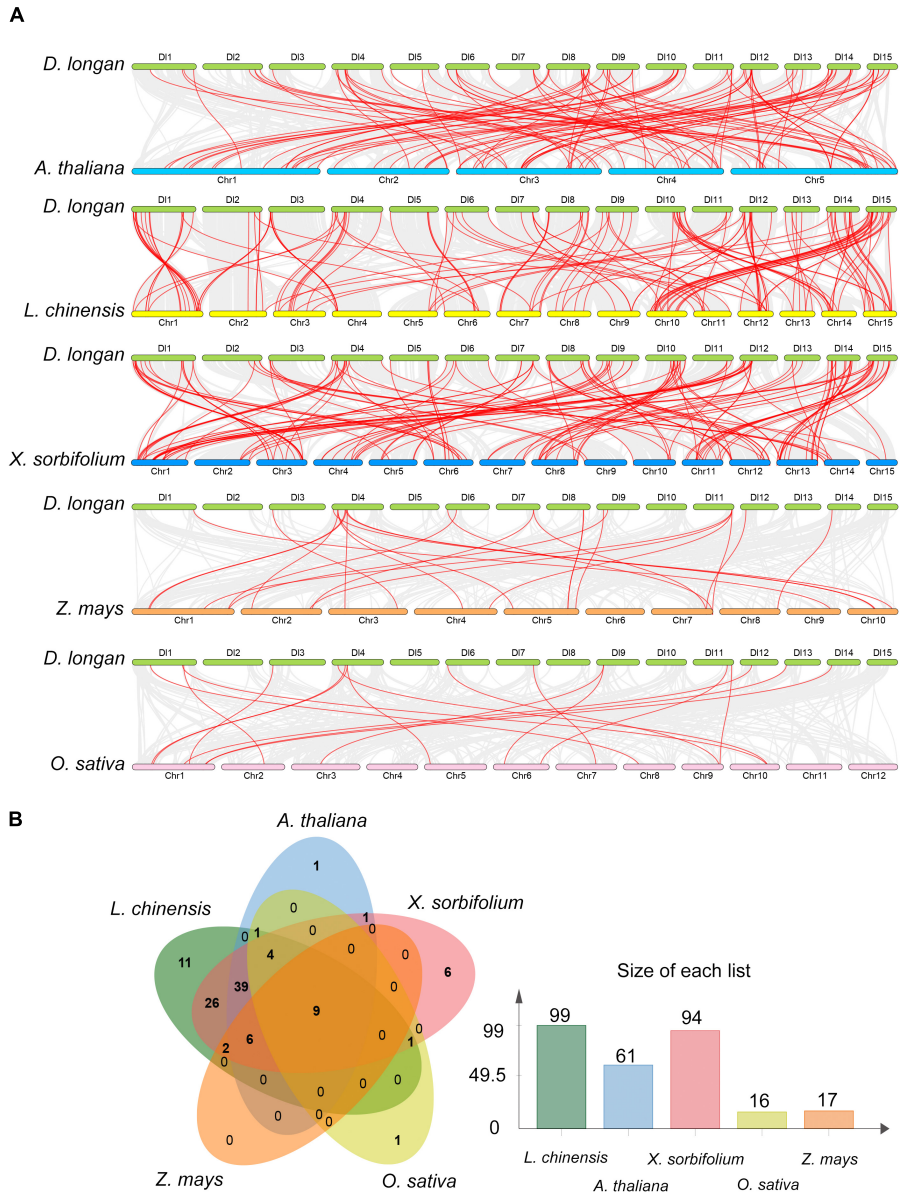


FIGURE 5 | Collinearity analysis of the *DlMYB* genes. **(A)** Synteny relationship of *R2R3MYB* genes between longan and five other representative plant species: Arabidopsis, litchi, yellow horn, maize, and rice. Gray lines in the background indicate the collinear blocks within two genomes, and red lines highlight the syntenic MYB gene pairs. **(B)** The *R2R3MYB* genes formed the colinear pairs between longan and the other five species. Bar graphs indicate the number of genes in colinear pairs.

showed that DlMYB16, DlMYB72, and DlMYB116 proteins were localized to the nucleus (**Figure 6**), which was consistent with subcellular localization predictions.

Expression Profile of *DlMYB* Genes in Different Tissues and Gene Ontology Annotation Analysis

The differential expression patterns of *DlR2R3-MYB* genes in nine tissues were analyzed using RNA-seq data (**Figure 7A**, **Supplementary Figures 3, 4**, and **Supplementary Tables 11, 12**). The results revealed that the *DlR2R3-MYB* genes showed varying expression patterns in different tissues and most genes displayed tissue-specific expressions. In general, 119 *R2R3-MYB* genes were clustered into seven groups according to the preferential expression tissues, which were designated as group I to group VII (**Figure 7A**). Group I genes were highly expressed in flowers and flower buds. Group II genes were widely expressed in different tissues, which indicated that they were involved in various organ formation processes in longan. Group III genes were mainly expressed in the roots, suggesting that they might execute certain functions in the root development. Genes

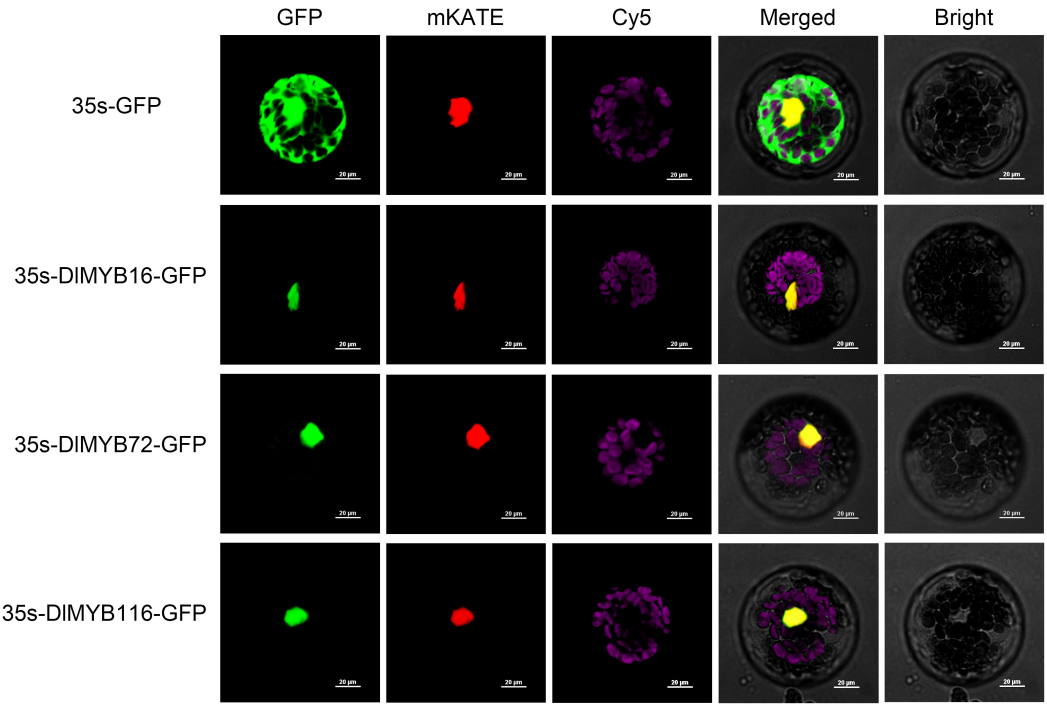


FIGURE 6 | Nuclear localization of DlMYB16, DlMYB72, and DlMYB116 proteins in Arabidopsis protoplasts. GFP: green fluorescent protein; mKATE: nuclear marker; Cy5: chloroplast; Bright: bright field; merged: combined mKATE and GFP and bright field. Bars, 20 μm.

in group IV-VII exhibited preferential expression in seeds, young fruits, pericarps, and pulps, respectively. In total, 74 (62.18%), 58 (48.74%), 47 (39.50%), 62 (52.10%), 57 (47.90%), 69 (57.98%), 30 (25.21%), 62 (52.10%), 50 (42.02%) *DlR2R3-MYB* genes were expressed (TMM > 1) in roots, stems, leaves, flower buds, flowers, young fruits, pulps, pericarps and seeds, respectively (**Supplementary Figure 3** and **Supplementary Table 12**). Flowers and flower buds were observed to have the most *DlR2R3-MYB* genes that had TMM ≥ 100, with 10.08 and 7.56%, respectively, suggesting that more *R2R3-MYB* genes were highly expressed in flower tissues than in other tissues. Of 119 *DlMYB* genes, 98 genes (82.35%) were expressed in at least one tissue (**Supplementary Table 11**). By contrast, five genes (*DlMYB11*, *DlMYB21*, *DlMYB38*, *DlMYB43*, and *DlMYB54*) in longan were not expressed (TMM = 0) in any of the examined tissues, possibly because they could only be expressed under given conditions, or detected in certain tissues that were not examined in our study, or they were pseudogenes. Three *DlMYB* genes (*DlMYB67*, *DlMYB70*, and *DlMYB79*) were observed to have high expression levels (TMM > 10) in all nine tissues examined. It is worth noting that fourteen genes (*DlMYB8*, *DlMYB16*, *DlMYB17*, *DlMYB33*, *DlMYB35*, *DlMYB67*, *DlMYB70*, *DlMYB72*, *DlMYB78*, *DlMYB79*, *DlMYB87*, *DlMYB93*, *DlMYB115*, and *DlMYB116*) showed extremely high expression levels (TMM > 100) in flowers and flower buds. Some genes exhibited flower-specific expressions. Four genes (*DlMYB28*, *DlMYB29*, *DlMYB42*, and *DlMYB109*) were expressed exclusively in flower buds, and *DlMYB87* was expressed only in flowers and flower buds.

A total of 153 GO items were enriched by the GO annotation with a cut-off value of $P \leq 0.05$, among which 146 GO terms were classified into the biological process category and only seven into the molecular function category (**Supplementary Table 13**). Fifty-eight of all *DlMYB* genes (48.74%) were assigned to categories related to hormones, stress response, and tissue development (**Figure 7B**). Interestingly, six DlMYB genes (*DlMYB19*, *DlMYB20*, *DlMYB30*, *DlMYB40*, *DlMYB72*, and *DlMYB116*) were assigned to the "flower development" category. Among them, *DlMYB19* belonged exclusively to the "flower development" category and *DlMYB20* was only associated with two categories "flower development" and "cellular response to hormone stimulus," while *DlMYB40* was also related to other several categories such as development and salt stress response (**Figure 7B** and **Supplementary Table 13**). Regarding their expression levels, however, *DlMYB40* was lowly expressed in all examined tissues except for roots and stems. *DlMYB19* and *DlMYB20* showed slightly higher expression (with TMM of 10∼12) in flower buds. Two genes, *DlMYB72* and *DlMYB116*, that had extremely high expression levels in flowers or flower buds were given more attention in the following analysis.

Expression Pattern of DlMYBs in Off-Season FI Induction by $KClO_3$

To further investigate the biological function of *DlMYB* genes responding to $KClO_3$ treatment, ten time points with untreated control and $KClO_3$ treatment were used to detect the expression profiles of *DlMYB* genes in apical buds. The hierarchical

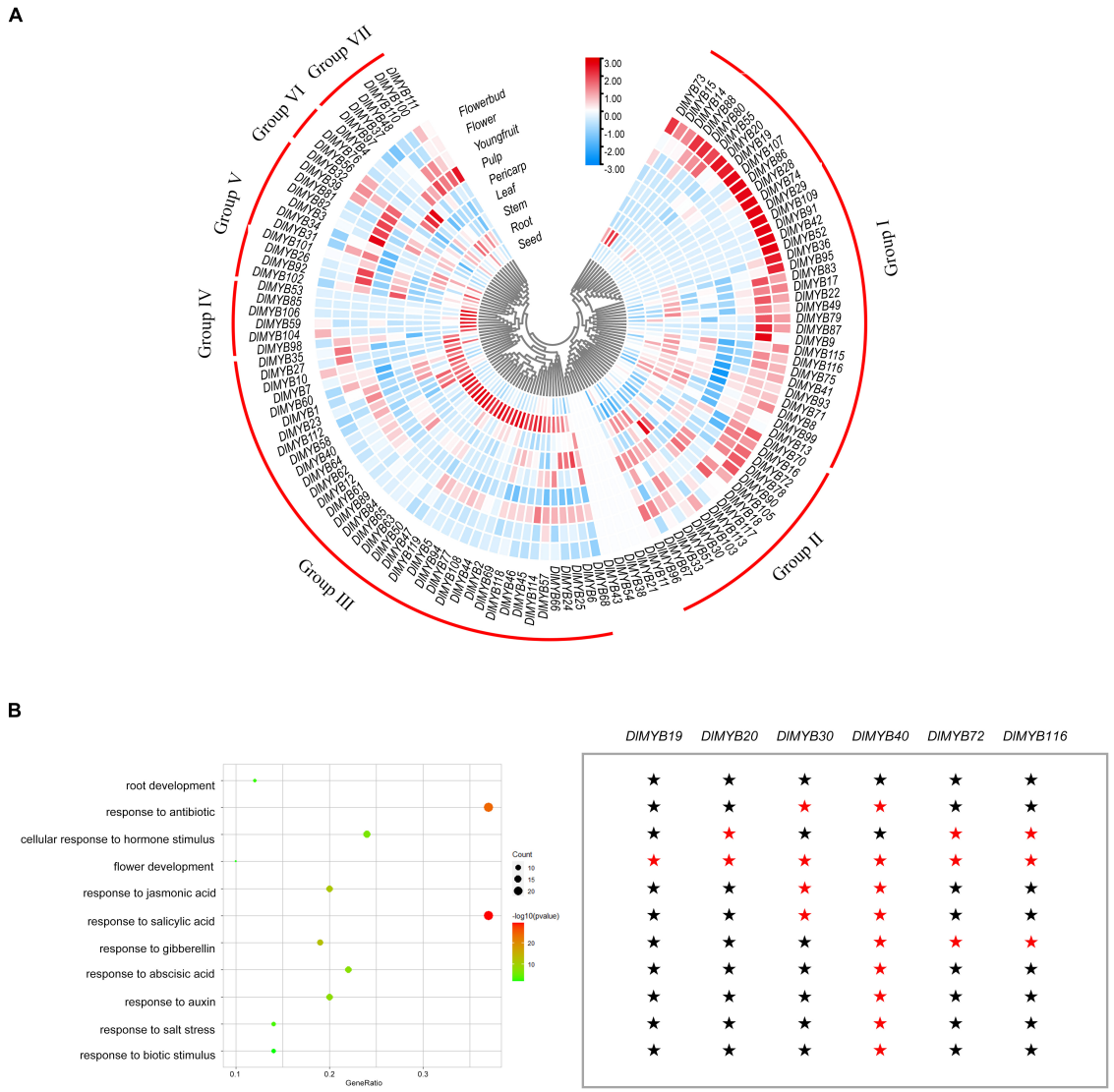


FIGURE 7 | Tissue-specific expression and GO enrichment analysis of DlMYB genes. **(A)** The expression levels of *DlMYB* genes in nine tissues were classified into groups I–VII. Different colors indicate log2 transformed expression levels, with red indicating high expression levels and blue indicating low expression levels. **(B)** Some *DlMYB* genes were classified as related to stress response, development, and hormones. The size of the circle indicates the number of *DlMYB* genes, and the color gradient indicates the size of the *P*-value. Six *DlMYB* genes are categorized into "flower development" function. A red pentagram represents that the gene is relevant, while a black pentagram indicates that the gene is not involved in the category.

cluster analysis was performed based on the expression (TMM values) of *DlMYB* genes treated with $KClO_3$ at different periods (**Figure 8A** and **Supplementary Table 14**). In the TMM analysis results, 20.17% (24/119) of all *DlMYB* genes were observed to have no expression (TMM < 1) at different time points (**Supplementary Table 14**). A total of 18 *DlMYB* genes (15.13%) were highly expressed (TMM > 10) at all ten time points, among which, *DlMYB35*, *DlMYB93*, *DlMYB98*, and *DlMYB101* were extremely highly expressed (TMM > 100). The time course analysis showed that six genes were significantly upregulated or downregulated (*Padj* < 0.05) under $KClO_3$ treatments at different time points compared with control. For instance, the gene expression of *DlMYB31* and *DlMYB68* were inhibited after $KClO_3$ treatment compared to the control, indicating they were negatively regulated by $KClO_3$ (**Figure 8B** and **Supplementary Table 14**). *DlMYB22* and *DlMYB71* were inhibited in the later stages after $KClO_3$ treatment. In contrast, the expression of *DlMYB23* and *DlMYB99* were promoted in the initial stages by the $KClO_3$ treatment. In addition, Some genes in the same subfamilies tended to exhibit similar expression patterns. For example, *DlMYB33* and *DlMYB116* in subfamily 20 were up-regulated at day 5 and day 10, and then commenced to decline at later time points (**Figure 2** and **Supplementary Table 14**).

To elucidate the regulatory pathway of off-season flowering among *DlMYBs* and other genes, coexpression networks were generated incorporating 16 modules via WGCNA. The heatmap

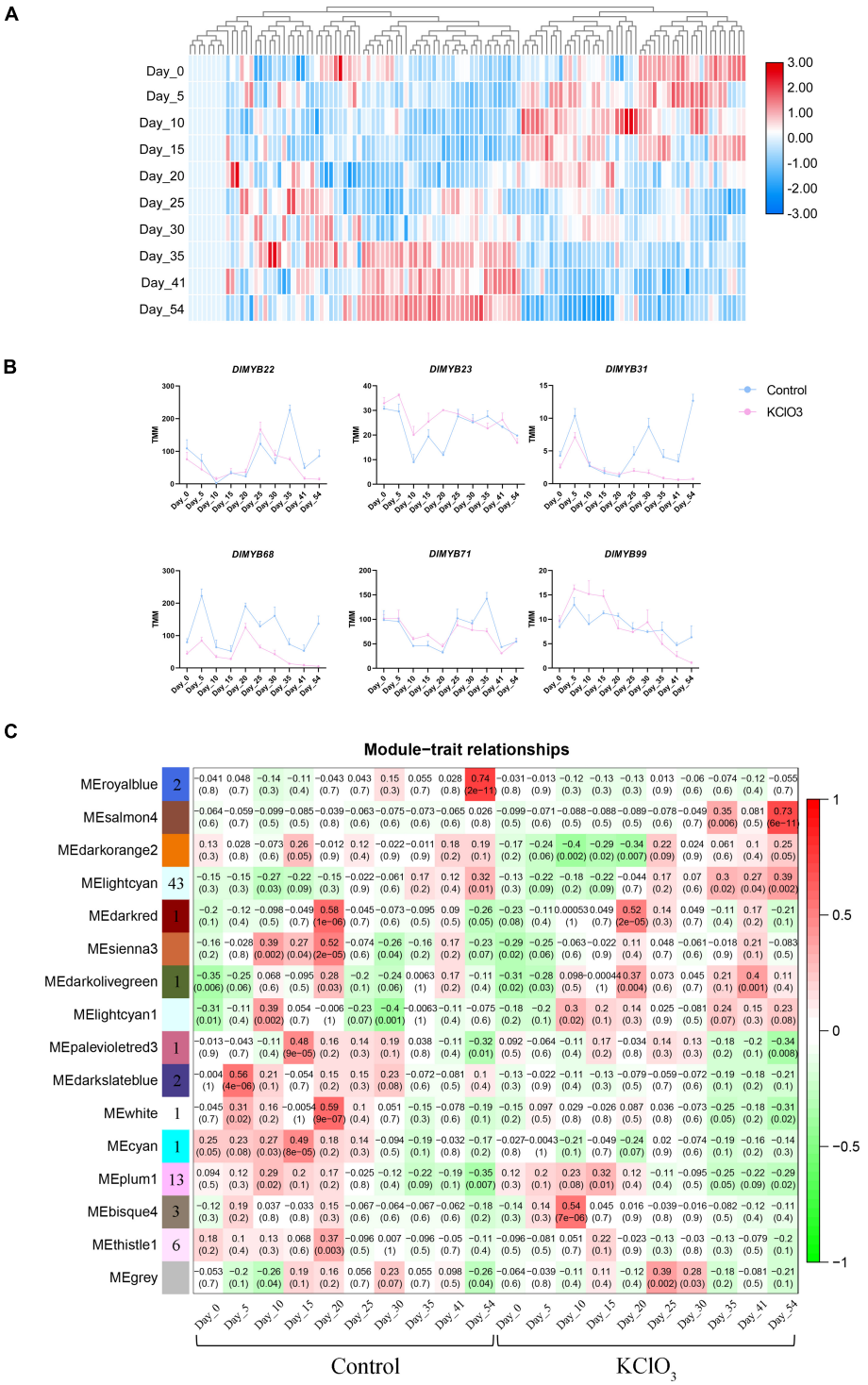


FIGURE 8 | Expression patterns of DlMYB genes under $KClO_3$ treatment. **(A)** The heat map displays the hierarchical clustering of the 119 *DlMYB* genes at different time points, with red indicating high expression levels and blue indicating low expression. **(B)** The line graphs display the effect of $KClO_3$ treatment on the expression of *DlMYB* genes. The blue line represents the control and the pink line represents $KClO_3$ treatment. **(C)** The heatmap from WGCNA of differentially expressed genes (DEGs) in off-season flower induction at different treatments and developmental stages. Each row corresponds to one module resulting from an WGCNA eigengene value. Each column corresponds to the control treatment and the $KClO_3$ treatment at different time-points. The color and number of each cell at the row-column intersection indicate the correlation coefficient between one specific module and one time point under control or $KClO_3$ treatment. Darker red represents a higher degree of correlation. The number in brackets indicate the significance (*p*-value) of the correlation coefficient. The number of *DlMYB* genes in each module is shown in the leftmost column.

of ten time points under control and $KClO_3$ treatment showed that 74 differentially expressed *R2R3-MYB* genes were distributed in eleven modules represented by different colors (MEbisque4, MEcyan, MEdarkolivegreen, MEdarkred, MEdarkslateblue, MElightcyan1, MEpalevioletred3, MEplum1, MEroyalblue, MEthistle1, and MEwhite) (**Figure 8C**). After filtering out the low-interacting genes based on the weight values reflecting the connectivity between genes in each module (the thresholds of weight values were shown in **Supplementary Table 15**), 35 *R2R3-MYB* genes remained that potentially interact with 567 genes. KEGG pathway enrichment analysis revealed that the "plant hormone signal transduction" pathway (ko04075) was enriched with 18 genes (**Supplementary Figure 9A**), and GO enrichment showed that these genes were mainly associated with "DNA binding TF activity" (**Supplementary Figure 9B**).

Among the eleven distinct modules, the MElightcyan module with 211 genes had an expression pattern correlated with the late period of $KClO_3$ (35-54 DAT) (**Figure 8C** and **Supplementary Table 15**). A gene expression network was developed to further mine candidate *DlMYB* genes (**Supplementary Figure 8**). Fifteen *DlMYB* genes were contained in the 211 genes. *DlMYB113* was identified as hub genes, with several TF genes such as *HISTONE H2A 11* (*HTA11, Dil.04g007090.1*), *EARLY IN SHORT DAYS 7/TILTED 1* (*TIL1, Dil.13g008690.1*), and *MITOTIC ARREST-DEFICIENT 2* (*MAD2*, *Dil.10g018750.1*). The MEplum1 module was found to be correlated with the early stage (15 DAT) of $KClO_3$ treatment. In total, 133 genes were involved in this module, including *DlMYB30*, *DlMYB41* and *DlMYB60*, and other genes such as *TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL FACTOR1-20* (*TCP20, Dil.12g008890.1*) and *HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 2* (*AHP2, Dil.04g009050.1*). The MEdarkolivegreen module with 7 genes were associated with the middle stage (20 DAT) of $KClO_3$ treatment. *DlMYB13* was identified as a hub gene in the central location, closely linked to *HUMAN WDR5 HOMOLOG A* (*WDR5a, Dil.04g005430.1*). The MEbisque4 module was comprised of 58 genes that were specifically and highly associated with the early period (10 DAT) of $KClO_3$ treatment. This module included three R2R3-MYB TF genes (*DlMYB8*, *DlMYB33*, and *DlMYB78*). *DlMYB*8 was identified as the top hub gene and was directly connected to *WRKY75* and *ERF114*.

Validation of *DlMYB* Genes Expression at Different Tissues

To validate the expression of *DlMYB* genes in flower development, eight representative *DlMYB* genes (*DlMYB16*, *DlMYB17*, *DlMYB19*, *DlMYB20*, *DlMYB35*, *DlMYB72*, *DlMYB78*, and *DlMYB116*) were selected for qRT-PCR analysis (**Figure 9**). Overall, the trends of the qRT-PCRs for the eight *DlMYB* genes were consistent with the results of RNA-seq analysis. However, a slight difference existed due to differences between longan cultivars used, the cultivation conditions of plants, and the geographical location of tissue collection.

As shown in **Figure 9**, *DlMYB17* and *DlMYB72* were observed to be expressed in the pericarps, flowers, and flower buds specifically. *DlMYB16* and *DlMYB116* were detected to have similar expression patterns and high expression in flowers and flower buds, suggesting a potential redundancy in their functions. Similar trends were found for *DlMYB19* and *DlMYB20* in subfamily 26. In addition, *DlMYB35* was specifically expressed in leaves, flowers, and flower buds. Strongly in line with our theorized expectations, the eight selected genes were significantly highly expressed in both flowers and flower buds compared with other tissues in longan.

DISCUSSION

The MYB gene family constitutes one of the largest TFs implicated in the regulation of a wide array of physiological processes in plants. Thus far, this present study is the first to conduct *in silico* systematic and genome-wide characterization of the *R2R3-MYB* gene family in the longan genome and reveal their potential roles in longan flower development.

Evolutionary Tracing of *DlMYB* Genes

In this study, a total of 119 R2R3-MYB proteins were identified in the "Shixia" longan genome, which were classified into 28 subgroups (**Figure 2** and **Supplementary Table 1**). A considerable fraction of subgroups (S1–S25) in our classification were consistent with the subgroupings in other plants (Dubos et al., 2010). Division of R2R3-MYBs into subgroups based on conserved sequence motifs would identify clades whose members commonly exhibited similar biological functions (Millard et al., 2019a). The numbers of *R2R3-MYB* genes in longan and Arabidopsis were remarkedly different in some subgroups within the phylogenetic tree (**Figure 2**). This phenomenon was also observed in many other species such as potato (Sun et al., 2019), maize (Du et al., 2012a), grape (Wong et al., 2016), and flax (Tombuloglu, 2020). For one reason, the number of *R2R3-MYB* genes may be related to the quality of the assembled genome. The marked improvement of the chromosome-level longan genome assembly would expand and enrich the predicted *DlMYB* gene pool. In comparison with our result, only 35 (Zheng et al., 2020) and 98 (Lin et al., 2017) *DlMYB* genes were identified based on transcriptome data and the previous scaffold-level longan assembly using Illumina data, respectively. For another, the variation in the total number of *R2R3-MYB* genes and subgroup genes among plants may have arisen from evolutionary divergence during which the expansion/contraction of *MYB* genes occurred in different species (Jiang and Rao, 2020). TD, SD and whole genome duplication (WGD) events have long been known to occur throughout plant evolution and perform a significant role in new gene generation and gene family expansion (Van de Peer, 2004). The results showed that tandem duplication and segmental duplication events were unevenly distributed throughout all chromosomes, with 16 *DlMYB* genes identified as TDs and nearly three times more *DlMYB* genes (44) classified as SDs (**Figure 4** and **Supplementary Tables 7, 8**). Similar trends

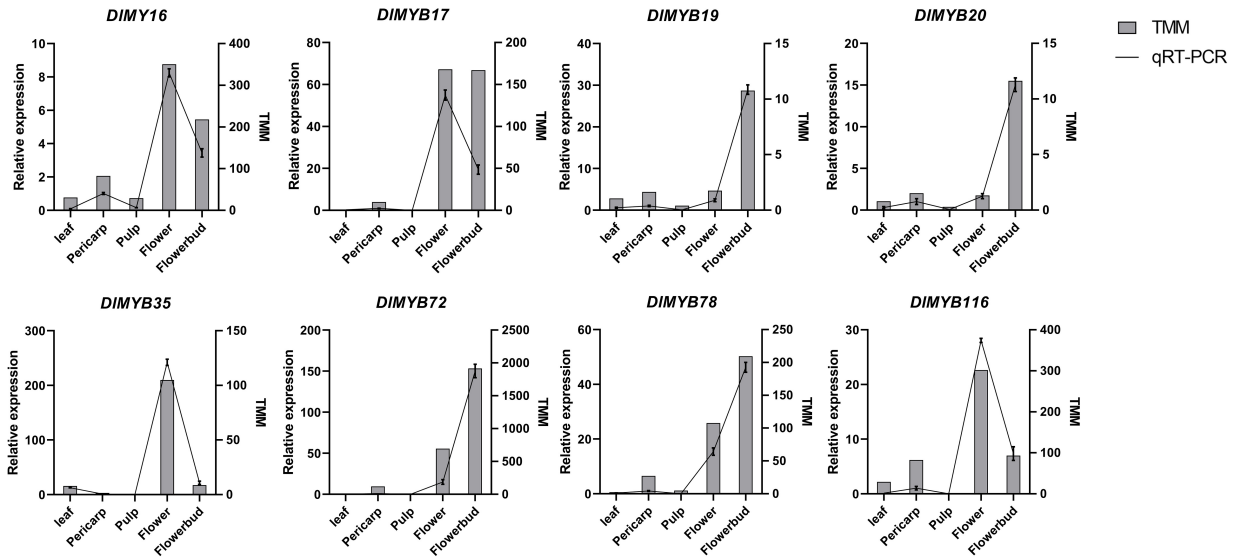


FIGURE 9 | qRT-PCR and RNA-seq analyses of eight selected DlMYB genes. The samples from flowers, flower buds, leaves, pulps, and pericarps show a range of expression changes of selected *DlMYB* genes. The left and right sides of the *Y*-axis indicate relative expression levels and TMM, respectively. To confirm the tissue-specific expression of the *DlMYB* genes, the expression pattern of selected *DlMYB* genes was calculated as a fold relative to the expression level of the pericarp.

of more SDs were observed in Arabidopsis (Cannon et al., 2004) and potato (Sun et al., 2019), suggesting a more crucial role of SD events in the expansion of *MYB* genes compared to TD events. More SD events and fewer TDs have commonly been present in gene families involved in housekeeping or core regulatory functions (Cannon et al., 2004). *Ka*/*Ks* was used to reflect the selection pressure on plants during evolution. The *Ka*/*Ks* ratios of all 35 tandemly and segmentally duplicated *DlR2R3-MYB* gene pairs were less than 1 (**Supplementary Table 8**), indicating that they might have been subject to purifying selection with constraint functional divergence after species-specific duplication events. This result corresponds well with previous observations in many plant species (Abbas et al., 2021; Wang et al., 2021). Although R2R3-MYB TFs contain "non-MYB regions" that are highly variable, disordered and compositionally biased, they have been evolving at a neutral rate and under the pressure of purifying selection with evolutionary constraints (Millard et al., 2019a).

Interspecies synteny analysis demonstrated that *DlR2R3-MYB* genes had higher collinearity with *R2R3-MYB* genes of three dicot species, especially Litchi, and yellow horn in the same family of *Sapindaceae* (**Figures 5A,B** and **Supplementary Tables 9, 10**), suggesting their divergence from the common dicot ancestor. In addition, some syntenic gene pairs were only present between longan and other dicots. For instance, 39 syntenic gene pairs were identified between longan and all examined dicots. Those pairs may appear after the divergence of dicotyledon and monocotyledon species. There were 26 syntenic *DlMYB* genes exclusively found in litchi and yellow horn, implying those genes may be *Sapindaceae* specific *MYB* genes, which have arisen after the divergence of *Sapindaceae* and *Brassicaceae*. Eighteen *AtMYB* genes paired with at least two longan genes indicate that additional longan genes would be generated due to duplication events and play vital roles during the evolution of longan. Nine *DlMYB* genes were identified to have synteny with their orthologs in five plants analyzed, suggesting that these genes may have existed prior to the divergence of these plant species and are highly conserved in evolution (**Figure 5B** and **Supplementary Tables 10, 11**).

Structural Characteristics and Functional Prediction of *DlMYB* Genes

We constructed a separate ML phylogenetic tree of 119 DlMYB proteins, which showed that either tandem replication or segmental replication gene pairs were classified into the same subgroup, except for the *DlMYB42-DlMYB79* segmental replication gene pair (**Figure 3A**). In addition, the motif composition and the distribution of exon/intron structures varied among *DlMYB* genes (**Figures 3B,C**). For example, *DlMYB98*, *DlMYB100*, and DlMYB101 contained only three motifs, while 16 *DlMYB* genes harbored the largest number of 7 motifs (**Figure 3B**). In general, a high degree of conservation in the MYB structural domain (R2 and R3 repeats) were found and members of the same subgroup tended to have similar motif combinations. The motifs that made up R2 and R3 repetitive sequences were highly conserved and mainly distributed at the N terminus, while specific motifs were present at the C terminus (**Figure 1** and **Supplementary Figure 1**). The motifs 9 and 17 belonged exclusively to S6, motifs 8 and 13 were only found in S27, and motif 15 was unique in S25. Previous researches revealed that unique motifs may be associated with the functional divergence of DlMYBs (Millard et al., 2019b). Gene structure analysis showed that members in the same subgroup

were also inclined to have similar exon/intron structures and intron numbers. Four genes in subgroup 6, *DlMYB52*, *DlMYB53*, *DlMYB54*, and *DlMYB55*, however, showed significantly different gene structures from each other although they formed a tandem gene cluster in Chr8. Consistent with previous studies (Abbas et al., 2021), most *R2R3-MYB* genes were found to contain only two introns with the exception of *DlMYB75* and *DlMYB55* that contained 13 and 8 introns, respectively. Introns were found to have functions in promoting cellular resistance to starvation by enhancing the repression of ribosomal protein genes (Parenteau et al., 2019).

The *cis*-acting element analysis results showed that DlMYB proteins were predicted to have various elements related to phytohormone, environmental stress, and cellular development (**Figure 3D** and **Supplementary Tables 5, 6**). Phytohormones such as ABA, ethylene, GA, and MeJA are involved in the regulation of plant flower development and senescence (Bao et al., 2020; Huang et al., 2021). The promoters of the *DlMYB* genes were predicted to contain several *cis*-acting elements associated with hormone responses, implying the potential role of *DlMYB* genes in the regulation of flower development. GO enrichment analysis also indicated that some DlMYB proteins have functions in response to hormones, tissue development and environmental stress (**Figure 7B** and **Supplementary Table 13**). Among them, the "flower development" pathway was enriched with six genes, indicating their potential functions in flowering (**Figure 7B** and **Supplementary Table 13**). In agreement with previous research results in Arabidopsis (Dubos et al., 2010), the MYB binding sites involved in flavonoid biosynthetic regulation (MBSI) elements were found in promoters of thirteen longan *R2R3-MYB* genes. Another 61 *DlMYB* genes contained the drought inducibility element known for its role in response to drought stress, suggesting that more than half of the *DlR2R3-MYB* genes might be involved in the drought stress response.

Expression Patterns and Potential Functions of *DlMYB* Genes Involved in Flowering

Many R2R3-MYB members are involved in regulating plant development (Millar and Gubler, 2005; Cheng et al., 2009; Phan et al., 2011; Song et al., 2011; Huang H. et al., 2017). The expression analysis allowed us to further estimate the functions of *DlMYB* genes in longan. Our results revealed that varying expression patterns were observed among longan tissues and most genes exhibited tissue-preferential expression (**Figure 7A**), suggesting their tissue-specific roles in the development of differentiated tissues. Five genes were not expressed in all examined tissues, which were probably pseudogenes or could only be detected in other tissues that we did not collect. The functions of *DlMYB* genes were closely linked to their orthologous counterparts in Arabidopsis. Four longan genes, *DlMYB18*, *DlMYB80*, *DlMYB81*, and *DlMYB90* clustered with *AtMYB33* and *AtMYB65* in subgroup 18, which could promote anther and pollen development in Arabidopsis (Millar and Gubler, 2005). In addition, *DlMYB18* and *DlMYB80* formed segmental replication gene pairs in the longan genome and had synteny with *AtMYB97*/*AtMYB101*/*AtMYB120* that were found to be synergistic pollen tube factors and could control the fertilization process (Liang et al., 2013). These four longan *MYB* genes abovementioned were found to be active in flowers or young fruits, suggesting their potential roles in pollen development. Similarly, *DlMYB19*, *DlMYB20, DlMYB96*, and *DlMYB109*, which grouped with *AtMYB80* in subgroup 26, also showed high expression in flowers, indicating they may participate in flower development, especially in regulating pollen development (Phan et al., 2011).

Moreover, *AtMYB21* and *AtMYB24* that take part in the regulation of anther development (Cheng et al., 2009; Song et al., 2011) grouped with *DlMYB87* in subgroup 19. *DlMYB87* was extremely highly expressed in flowers, suggesting its possible involvement in regulating anther development. *AtMYB108* in subfamily 20 can regulate stamen maturation in Arabidopsis (Mandaokar and Browse, 2009). Another gene *AtMYB62* in S20 is a regulator of phosphorus stress in Arabidopsis. Overexpression of *AtMYB62* inhibits the synthesis of gibberellic acid (GA) and the expression of the flowering regulators *SCO1* and *SUPERMAN*, ultimately leading to delayed flowering in Arabidopsis (Devaiah et al., 2009). Their homologs in longan, *DlMYB16*, *DlMYB33*, *DlMYB72*, and *DlMYB116* were observed to have high expression in flowers. Among them, *DlMYB72* and *DlMYB116* had similar expression patterns as segmental replication genes and formed syntenic gene pairs with *AtMYB62*/*AtMYB116* and *AtMYB62*, respectively (**Figure 7A** and **Supplementary Tables 10, 11**). Therefore, these genes may have similar functional characteristics in the regulation of flower development. Beyond its function in flowering, *AtMYB108* also participated in the regulation of the leaf senescence process, acting by binding to a specific region of *ANAC003* promoter to form a MYB-NAC regulatory complex (Chou et al., 2018). Its homologs *DlMYB16* and *DlMYB116* were also highly expressed in leaves, suggesting that genes in the same subgroup may play different biological functions. In subgroup 14, *AtMYB68* was demonstrated to specifically regulate the root growth (Jiang et al., 2004), and its homologs *DlMYB8*, *DlMYB12*, *DlMYB44*, *DlMYB66*, *DlMYB99*, and *DlMYB119* were also found to be highly expressed in the roots, presumably having an important role in longan root development. The high levels of association between longan MYB expression patterns and the roles of their orthologs in Arabidopsis may be evidence of functional conservation between homologous R2R3-MYB members of Arabidopsis and longan.

The qRT-PCR experiments demonstrated the high expression levels of *DlMYB19*, *DlMYB20*, *DlMYB72*, and *DlMYB116* in flowers and flower buds, which were also annotated in the "flower development" pathway by GO enrichment analysis, further suggesting that these genes may play an active part in the regulation of flower development (**Figure 9**).

Potential Functions of *DlMYB* Genes Related to Stress and Metabolite Synthesis

Multiple R2R3MYB proteins were found to be capable of abiotic and biotic stress tolerance in plants (Dubos et al., 2010).

AtMYB15, which clustered with the segmental replication gene pairs *DlMYB13* and *DlMYB31* in subgroup 2, was demonstrated to possess cold stress tolerance and drought resistance (Ding et al., 2009; Kim et al., 2017). *AtMYB15* enhances the susceptibility to freezing by negatively regulating *CBF* genes which are essential for freezing resistance in Arabidopsis, whereas MPK6-mediated phosphorylation reduces the affinity of *AtMYB15* to bind to the *CBF3* promoter to enhance cold stress tolerance (Kim et al., 2017). Meanwhile, the overexpression of *AtMYB15* enhanced ABA-mediated drought resistance and salt tolerance in Arabidopsis (Ding et al., 2009). Interestingly, *DlMYB13* and *DlMYB31* happened to have low-temperature responsiveness (LTR) elements and several ABA-responsive (ABRE) elements, and *DlMYB31* was assigned to the "response to cold" pathway in GO enrichment analysis (**Supplementary Tables 5, 6, 9–13**), indicating their potential role in frost and drought resistance. Also, lignin biosynthesis is regulated by *AtMYB15*, which plays a key role in regulating plant growth and development as well as in plant innate immunity by promoting the synthesis of secondary cell walls (Kim et al., 2020). These two homologous *DlMYB* genes may also have similar biological functions that remain to be uncovered. In subgroup 22, five longan genes (*DlMYB36*, *DlMYB43*, *DlMYB67*, *DlMYB68*, and *DlMYB78*) formed a single cluster with four orthologous *AtMYB* genes (*AtMYB44*, *AtMYB70*, *AtMYB73*, and *AtMYB77*), which have been reported to be involved in the abiotic stress response (Jung et al., 2008; Zhao et al., 2014; Yang et al., 2020). For example, *AtMYB44* among them functions in the regulation of ABA-mediated stomatal closure as well as drought tolerance (Jung et al., 2008). The *cis*-acting element prediction showed that these four longan genes all have drought inducibility elements except for *DlMYB68* (**Supplementary Tables 5, 6**). In particularly, *DlMYB78* contained four drought inducibility elements, indicating it may have greater potential in response to drought induction. In terms of transcript expression, *DlMYB67* was extremely highly expressed in all tissues. qRT-PCR experiments showed that *DlMYB78* was extremely highly expressed in flowers and flower buds, which may promote flowering under drought induction (**Figure 9** and **Supplementary Tables 9–11**).

R2R3-MYB proteins are involved in different primary and secondary metabolic regulations in plants (Dubos et al., 2010). In Arabidopsis subgroup5, *AtMYB123* ensures a high accumulation of proanthocyanidin (PAs) by forming the MYB-BHLH-WDR (MBW) complex (Baudry et al., 2004). The AtMYB123 homolog, *DlMYB7*, was highly expressed in roots, stems, leaves, flowers and young fruits, with a possible function in proanthocyanidin synthesis (**Figure 7A** and **Supplementary Figure 4**). Four *AtMYB* genes, *AtMYB75*/*PAP1*, *AtMYB90*/*PAP2*, *AtMYB113*, and *AtMYB114* have been reported to control anthocyanin biosynthesis (Teng et al., 2005; Gonzalez et al., 2008), which formed a subgroup (S6) with the tandem replication gene cluster *DlMYB52*/*DlMYB53*/*DlMYB54*/*DlMYB55* in longan (**Figures 2**, **4A**). *DlMYB32*, *AtMYB11*/*PFG2*, *ATMYB12*/*PFG1*, and *AtMYB111*/*PFG3* were classified as subgroup 7. The *AtMYB*s in this group mainly participate in the regulation of flavonol biosynthesis (Dubos et al., 2010). Previous studies have found *AtMYB21* and *AtMYB24* in subgroup 19, that cluster with *DlMYB87*, played conserved regulatory roles in flavonol biosynthesis in Arabidopsis anthers and pollen by activating the *flavonol synthase 1* (*AtFLS 1*) gene (Stracke et al., 2010; Shan et al., 2020). The flower-specific expressed gene *DlMYB42* clustered with *AtMYB28* (*HAG1*), *AtMYB29* (*HAG2*), and *AtMYB86* (*HAG2*) with functions involved in aliphatic glucosinolate biosynthesis (Gigolashvili et al., 2008) in subgroup 12.

Expression of *DlMYB* Genes in Off-Season Induced Flowering by $KClO_3$

The transcriptome dynamics of 119 *DlMYB* genes in apical buds at different periods under $KClO_3$ treatment were analyzed(**Figure 8A** and **Supplementary Table 14**). The time course analysis of 119 *DlMYB* genes at different periods under control and $KClO_3$ treatment promoted the expression of *DlMYB23* and *DlMYB99* at the initial stage, which may be activated in response to the $KClO_3$ induction (**Figure 8B** and **Supplementary Table 14**). Compared with the gene expression under control without $KClO_3$ treatment, we observed that the expression of four genes (*DlMYB22*, *DlMYB31*, *DlMYB68*, and *DlMYB71*) were inhibited under $KClO_3$ treatment, indicating that they may have a negative effect on flowering by $KClO_3$ treatment. In general, *DlMYB* genes in the same subfamily tended to have similar expression patterns.

Co-expression network analysis via WGCNA was used to identify the role of *DlMYB* genes in $KClO_3$-mediated FI (**Figure 8C**). Based on the FLOR-ID of Arabidopsis (Bouché et al., 2016), we found seven flower-related genes out of these 567 genes, including *HTA11*, *WDR5a*, *TIL1*, *SENSITIVE TO FREEZING 6* (*SFR6*), *HISTONE ACETYLTRANSFERASE OF THE MYST FAMILY 1* (*HAM1*), *CURLY LEAF* (*CLF*), and *CYCLING DOF FACTOR 3* (*CDF3*). Some *DlMYB* hub genes were closely linked to other TF genes. For example, three *DlMYB* genes (*DlMYB8*, *DlMYB33*, and *DlMYB78*) were directly linked to TF genes such as *WRKY75* and *ERF114* in the MEbisque4 module (**Supplementary Figure 8**). The *WRKY75* could positively regulate flowering in Arabidopsis by GA-mediated binding to the promoter of the *FT* gene (Zhang et al., 2018). *WRKY2* and *WRKY34* repressed *AtMYB97*, *AtMYB101*, and *AtMYB120* expression during male gametogenesis (Lei et al., 2018). *DlMYB8* and *WRKY75* were up-regulated in the early stages of $KClO_3$ treatment and presumably they may positively coregulate *FT* gene expression and eventually induce flowering in longan. In the MEroyalblue module, on the contrary, *DlMYB31*, *DlMYB51* and *ERF1* were down-regulated after $KClO_3$ treatment. In Arabidopsis, *ERF1* delayed flowering by repressing the expression of the *FT* gene (Chen et al., 2021). The downregulation of those genes under $KClO_3$ may attenuate the repression of *FT* expression and ultimately promoted flowering. In the MEplum1 module, *DlMYB41*, *TCP20*, and *AHP2* were all up-regulated in the earlier stages after $KClO_3$ treatment (**Supplementary Figure 8** and **Supplementary Table 14**). In Arabidopsis, interestingly, *AHP2* functioned in the cytokinin pathway to promote female gametophyte development and *TCP20* promoted the absorption of nitrate (Guan et al., 2014; Liu et al., 2017).

Chlorate was a nitrate analog that could be absorbed by the nitrate transporters (Borges et al., 2004). Thus, *DlMYB41* and *TCP20* may together facilitate the utilization of $KClO_3$, which was subsequently reduced to chlorite and hypochlorite and directly induced the stress response. Meanwhile, *DlMYB41* and *AHP2* were interacted to stimulate CK-mediated signal transduction. Eventually, they acted together to influence the expression of flowering-related genes and induce flowering. Whether and how these *DlMYB* genes function in longan flowering induction and development requires further experiments to elucidate.

CONCLUSION

In the current study, a systematic identification of R2R3-MYB gene family members in longan was carried out. We identified a total of 119 full-length *R2R3-MYB* genes in longan, which were further divided into 28 distinct subgroups by phylogenetic classification. A comprehensive bioinformatics analysis was performed to investigate the protein physicochemical properties, motif composition, gene structure, and *cis*-element prediction of the *DlMYB* genes. An uneven distribution of *DlMYB* genes on 15 chromosomes was observed with at least two genes per chromosome. The synteny analysis indicated segmental replication events were major contributors to the expansion of the *R2R3-MYB* gene family in longan. Comparative syntenic mapping among longan and other five representative species provided insights into the evolutionary relationship of orthologous MYB members and implied the presence of functional differentiation led by duplication events during evolution. Subcellular localization analysis revealed that the selected flowering-associated *DlMYB* genes mainly function in the nucleus as predicted. Expression profiling of *DlMYB* genes indicates their tissue-preferential expression pattern and $KClO_3$-induced off-season flowering effect. GO enrichment analysis revealed that six *DlMYB* genes were assigned to the "flower development" category. The selected eight genes were further proved to be highly expressed in flowers and flower buds by qRT-PCR, indicating their possible involvement in the flower development process. Expression patterns in $KClO_3$-mediated FI indicated two up-regulated and four down-regulated *DlR2R3-MYB* genes were involved. These results will open up untapped avenues for further functional analysis of the *R2R3-MYB* genes to elucidate their roles in longan FI and floral organ development.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

JF and JZ conceived the study and designed the experiments. QC, JF, XZ, YF, BW, SX, and KZ carried out the experiments and analyzed the data. QC and JF wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the Natural Science Foundation of Fujian Province, China under Grant Number 2019J0102; the Natural Science Foundation of China under Grant Number 41906096; the Middle-aged Teachers from Fujian Provincial Department of Education Grant Number JT180082; and the Scientific Research Innovation Program "Xiyuanjiang River Scholarship" of College of Life Sciences, Fujian Normal University under Grant Number 22FSSK011.

ACKNOWLEDGMENTS

We would like to thank the editor and reviewers for their constructive comments on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.820439/full#supplementary-material

REFERENCES

Abbas, F., Ke, Y., Zhou, Y., Yu, Y., Waseem, M., Ashraf, U., et al. (2021). Genome-wide analysis reveals the potential role of MYB transcription factors in floral scent formation in *Hedychium coronarium*. *Front. Plant Sci.* 12:623742. doi: 10.3389/fpls.2021.623742

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335

Bao, S., Hua, C., Shen, L., and Yu, H. (2020). New insights into gibberellin signaling in regulating flowering in Arabidopsis. *J. Integr. Plant Biol.* 62, 118–131. doi: 10.1111/jipb.12892

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

Borges, R., Miguel, E. C., Dias, J. M. R., Da Cunha, M., Bressan-Smith, R. E., De Oliveira, J. G., et al. (2004). Ultrastructural, physiological and biochemical analyses of chlorate toxicity on rice seedlings. *Plant Sci.* 166, 1057–1062. doi: 10.1016/j.plantsci.2003.12.023

Bouché, F., Lobet, G., Tocquin, P., and Périlleux, C. (2016). FLOR-ID: an interactive database of flowering-time gene networks in *Arabidopsis thaliana*. *Nucleic Acids Res.* 44, D1167–D1171. doi: 10.1093/nar/gkv1054

Braun, E. L., and Grotewold, E. (1999). Newly discovered plant c-myb-like genes rewrite the evolution of the plant myb gene family. *Plant Physiol.* 121, 21–24.

Briesemeister, S., Rahnenführer, J., and Kohlbacher, O. (2010). YLoc–an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 38, W497–W502. doi: 10.1093/nar/gkq477

Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 4:10. doi: 10.1186/1471-2229-4-10

Chang, D., and Duda, T. F. Jr. (2012). Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol.* 29, 2019–2029. doi: 10.1093/molbev/mss068

Chang, Q. (2010). *Study on Relationship between off-Season Floral Induction in Longan and Nutrition of the Carbon and Nitrogen*. Fuzhou: Fujian Agriculture & Forestry University. *Master's thesis*.

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chen, Y., Zhang, L., Zhang, H., Chen, L., and Yu, D. (2021). *ERF1* delays flowering through direct inhibition of *FLOWERING LOCUS T* expression in Arabidopsis. *J. Integr. Plant Biol.* 63, 1712–1723. doi: 10.1111/jipb.13144

Cheng, H., Song, S., Xiao, L., Soo, H. M., Cheng, Z., Xie, D., et al. (2009). Gibberellin acts through jasmonate to control the expression of *MYB21*, *MYB24*, and *MYB57* to promote stamen filament growth in Arabidopsis. *PLoS Genet.* 5:e1000440. doi: 10.1371/journal.pgen.1000440

Cho, L. H., Yoon, J., and An, G. (2017). The control of flowering time by environmental factors. *Plant J.* 90, 708–719. doi: 10.1111/tpj.13461

Chou, M. L., Liao, W. Y., Wei, W. C., Li, A. Y., Chu, C. Y., Wu, C. L., et al. (2018). The direct involvement of dark-Induced Tic55 protein in chlorophyll catabolism and its indirect role in the MYB108-NAC signaling pathway during leaf senescence in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 19:1854. doi: 10.3390/ijms19071854

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190. doi: 10.1101/gr.849004

Devaiah, B. N., Madhuvanthi, R., Karthikeyan, A. S., and Raghothama, K. G. (2009). Phosphate starvation responses and gibberellic acid biosynthesis are regulated by the *MYB62* transcription factor in Arabidopsis. *Mol. Plant* 2, 43–58. doi: 10.1093/mp/ssn081

Ding, Z., Li, S., An, X., Liu, X., Qin, H., and Wang, D. (2009). Transgenic expression of *MYB15* confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*. *J. Genet. Genomics* 36, 17–29. doi: 10.1016/s1673-8527(09)60003-5

Du, H., Yang, S. S., Liang, Z., Feng, B. R., Liu, L., Huang, Y. B., et al. (2012b). Genome-wide analysis of the MYB transcription factor superfamily in soybean. *BMC Plant Biol.* 12:106. doi: 10.1186/1471-2229-12-106

Du, H., Feng, B. R., Yang, S. S., Huang, Y. B., and Tang, Y. X. (2012a). The R2R3-MYB transcription factor gene family in maize. *PLoS One* 7:e37463. doi: 10.1371/journal.pone.0037463

Du, H., Liang, Z., Zhao, S., Nan, M. G., Tran, L. S., Lu, K., et al. (2015). The evolutionary history of R2R3-MYB proteins across 50 eukaryotes: new insights into subfamily classification and expansion. *Sci. Rep.* 5:11037. doi: 10.1038/srep11037

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci.* 15, 573–581. doi: 10.1016/j.tplants.2010.06.005

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563

Gigolashvili, T., Engqvist, M., Yatusevich, R., Müller, C., and Flügge, 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

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

Guan, P., Wang, R., Nacry, P., Breton, G., Kay, S. A., Pruneda-Paz, J. L., et al. (2014). Nitrate foraging by Arabidopsis roots is mediated by the transcription factor TCP20 through the systemic signaling pathway. *Proc. Natl. Acad. Sci. U. S. A.* 111, 15267–15272. doi: 10.1073/pnas.1411375111

Hong, L., Niu, F., Lin, Y., Wang, S., Chen, L., and Jiang, L. (2021). MYB106 is a negative regulator and a substrate for $CRL3^{BPM}$ E3 ligase in regulating flowering time in *Arabidopsis thaliana*. *J Integr. Plant Biol.* 63, 1104–1119. doi: 10.1111/jipb.13071

Huang, F. N., Fu, Z. Y., Zeng, L. H., and Morley-Bunker, M. (2017). Isolation and characterization of *GI* and *FKF1* homologous genes in the subtropical fruit tree *Dimocarpus longan*. *Mol. Breed.* 37:13. doi: 10.1007/s11032-017-0691-z

Huang, H., Gao, H., Liu, B., Qi, T., Tong, J., Xiao, L., et al. (2017). Arabidopsis *MYB24* regulates jasmonate-mediated stamen sevelopment. *Front. Plant Sci.* 8:1525. doi: 10.3389/fpls.2017.01525

Huang, S., Han, D., Wang, J., Guo, D., and Li, J. (2021). Floral induction of longan (*Dimocarpus longan*) by potassium chlorate: application, mechanism, and future perspectives. *Front. Plant Sci.* 12:670587. doi: 10.3389/fpls.2021.670587

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47, D309–D314. doi: 10.1093/nar/gky1085

Hurst, L. D. (2002). The Ka/Ks ratio: diagnosing the form of sequence evolution. *Trends Genet.* 18:486. doi: 10.1016/s0168-9525(02)02722-1

Jia, L., Clegg, M. T., and Jiang, T. (2004). Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies indica and japonica genomes. *Plant Physiol.* 134, 575–585. doi: 10.1104/pp.103.027201

Jia, T. Q., Wei, D. F., Meng, S., Allan, A. C., and Zeng, L. H. (2014). Identification of regulatory genes implicated in continuous flowering of longan (*Dimocarpus longan* L.). *PLoS One* 9:e114568. doi: 10.1371/journal.pone.0114568

Jiang, C., Gu, J., Chopra, S., Gu, X., and Peterson, T. (2004). Ordered origin of the typical two- and three-repeat Myb genes. *Gene* 326, 13–22. doi: 10.1016/j.gene.2003.09.049

Jiang, C. K., and Rao, G. Y. (2020). Insights into the diversification and evolution of R2R3-MYB transcription factors in plants. *Plant Physiol.* 183, 637–655. doi: 10.1104/pp.19.01082

Jung, C., Seo, J. S., Han, S. W., Koo, Y. J., Kim, C. H., Song, S. I., et al. (2008). Overexpression of *AtMYB44* enhances stomatal closure to confer abiotic stress tolerance in transgenic Arabidopsis. *Plant Physiol.* 146, 623–635. doi: 10.1104/pp.107.110981

Kanei-Ishii, C., Sarai, A., Sawazaki, T., Nakagoshi, H., He, D. N., Ogata, K., et al. (1990). The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the myb protooncogene product. *J. Biol. Chem.* 265, 19990–19995. doi: 10.1016/S0021-9258(17)45472-X

Katiyar, A., Smita, S., Lenka, S. K., Rajwanshi, R., Chinnusamy, V., and Bansal, K. C. (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC Genomics* 13:544. doi: 10.1186/1471-2164-13-544

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kim, S. H., Kim, H. S., Bahk, S., An, J., Yoo, Y., Kim, J. Y., et al. (2017). Phosphorylation of the transcriptional repressor *MYB15* by mitogen-activated protein kinase 6 is required for freezing tolerance in Arabidopsis. *Nucleic Acids Res.* 45, 6613–6627. doi: 10.1093/nar/gkx417

Kim, S. H., Lam, P. Y., Lee, M. H., Jeon, H. S., Tobimatsu, Y., and Park, O. K. (2020). The Arabidopsis R2R3 MYB transcription factor *MYB15* is a key regulator of lignin biosynthesis in effector-triggered immunity. *Front. Plant Sci.* 11:583153. doi: 10.3389/fpls.2020.583153

Klempnauer, K. H., Gonda, T. J., and Bishop, J. M. (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. *Cell* 31, 453–463. doi: 10.1016/0092-8674(82)90138-6

Kranz, H., Scholz, K., and Weisshaar, B. (2000). c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *Plant J.* 21, 231–235. doi: 10.1046/j.1365-313x.2000.00666.x

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., Mcwilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Lee, H. G., and Seo, P. J. (2015). The MYB96-HHP module integrates cold and abscisic acid signaling to activate the CBF-COR pathway in Arabidopsis. *Plant J.* 82, 962–977. doi: 10.1111/tpj.12866

Lei, R., Ma, Z., and Yu, D. (2018). WRKY2/34-VQ20 Modules in *Arabidopsis thaliana* negatively regulate expression of a trio of related MYB transcription factors during pollen development. *Front. Plant Sci.* 9:331. doi: 10.3389/fpls.2018.00331

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, W., Liu, Y., Zhao, J., Zhen, X., Guo, C., and Shu, Y. (2019). Genome-wide identification and characterization of R2R3-MYB genes in *Medicago truncatula*. *Genet. Mol. Biol.* 42, 611–623. doi: 10.1590/1678-4685-gmb-2018-0235

Liang, Y., Tan, Z. M., Zhu, L., Niu, Q. K., Zhou, J. J., Li, M., et al. (2013). *MYB97*, *MYB101* and *MYB120* function as male factors that control pollen tube-synergid interaction in *Arabidopsis thaliana* fertilization. *PLoS Genet.* 9:e1003933. doi: 10.1371/journal.pgen.1003933

Lin, Y., Min, J., Lai, R., Wu, Z., Chen, Y., Yu, L., et al. (2017). Genome-wide sequencing of longan (*Dimocarpus longan* Lour.) provides insights into molecular basis of its polyphenol-rich characteristics. *Gigascience* 6, 1–14. doi: 10.1093/gigascience/gix023

Liu, Z., Yuan, L., Song, X., Yu, X., and Sundaresan, V. (2017). AHP2, AHP3, and AHP5 act downstream of CKI1 in Arabidopsis female gametophyte development. *J. Exp. Bot.* 68, 3365–3373. doi: 10.1093/jxb/erx181

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

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Mandaokar, A., and Browse, J. (2009). *MYB108* acts together with *MYB24* to regulate jasmonate-mediated stamen maturation in Arabidopsis. *Plant Physiol.* 149, 851–862. doi: 10.1104/pp.108.132597

Matsumoto, T. K. (2006). Genes uniquely expressed in vegetative and potassium chlorate induced floral buds of *Dimocarpus longan*. *Plant Sci.* 170, 500–510. doi: 10.1016/j.plantsci.2005.09.016

Millar, A. A., and Gubler, F. (2005). The Arabidopsis GAMYB-like genes, *MYB33* and *MYB65*, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* 17, 705–721. doi: 10.1105/tpc.104.027920

Millard, P. S., Kragelund, B. B., and Burow, M. (2019a). R2R3 MYB transcription factors - functions outside the DNA-binding domain. *Trends Plant Sci.* 24, 934–946. doi: 10.1016/j.tplants.2019.07.003

Millard, P. S., Weber, K., Kragelund, B. B., and Burow, M. (2019b). Specificity of MYB interactions relies on motifs in ordered and disordered contexts. *Nucleic Acids Res.* 47, 9592–9608. doi: 10.1093/nar/gkz691

Naik, J., Rajput, R., Pucker, B., Stracke, R., and Pandey, A. (2021). The R2R3-MYB transcription factor MtMYB134 orchestrates flavonol biosynthesis in *Medicago truncatula*. *Plant Mol. Biol.* 106, 157–172. doi: 10.1007/s11103-021-01135-x

Nguyen Ba, A. N., Pogoutse, A., Provart, N., and Moses, A. M. (2009). NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* 10:202. doi: 10.1186/1471-2105-10-202

Nicholas, K. B., Nicholas, H., and Deerfield, D. (1997). GeneDoc: analysis and visualization of genetic variation. *EMBNEW News* 4:14.

Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., et al. (1992). Solution structure of a DNA-binding unit of Myb: a helix-turn-helix-related motif with conserved tryptophans forming a hydrophobic core. *Proc. Natl. Acad. Sci. U. S. A.* 89, 6428–6432. doi: 10.1073/pnas.89.14.6428

Parenteau, J., Maignon, L., Berthoumieux, M., Catala, M., Gagnon, V., and Abou Elela, S. (2019). Introns are mediators of cell response to starvation. *Nature* 565, 612–617. doi: 10.1038/s41586-018-0859-7

Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A., and Saedler, H. (1987). The regulatory c1 locus of *Zea mays* encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *Embo J.* 6, 3553–3558. doi: 10.1002/j.1460-2075.1987.tb02684.x

Phan, H. A., Iacuone, S., Li, S. F., and Parish, R. W. (2011). The *MYB80* transcription factor is required for pollen development and the regulation of tapetal programmed cell death in *Arabidopsis thaliana*. *Plant Cell* 23, 2209–2224. doi: 10.1105/tpc.110.082651

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204. doi: 10.1093/nar/gky448

Pucker, B., Pandey, A., Weisshaar, B., and Stracke, R. (2020). The R2R3-MYB gene family in banana (*Musa acuminata*): genome-wide identification, classification and expression patterns. *PLoS One* 15:e0239275. doi: 10.1371/journal.pone.0239275

Rangkadilok, N., Worasuttayangkurn, L., Bennett, R. N., and Satayavivad, J. (2005). Identification and quantification of polyphenolic compounds in longan (*Euphoria longana* Lam.) fruit. *J. Agric. Food Chem.* 53, 1387–1392. doi: 10.1021/jf0403484

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

Rosinski, J. A., and Atchley, W. R. (1998). Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *J. Mol. Evol.* 46, 74–83. doi: 10.1007/pl00006285

Roy, S. (2016). Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. *Plant Signal. Behav.* 11:e1117723. doi: 10.1080/15592324.2015.1117723

Seo, P. J., Xiang, F., Qiao, M., Park, J. Y., Lee, Y. N., Kim, S. G., et al. (2009). The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in Arabidopsis. *Plant Physiol.* 151, 275–289. doi: 10.1104/pp.109.144220

Shan, X., Li, Y., Yang, S., Yang, Z., Qiu, M., Gao, R., et al. (2020). The spatio-temporal biosynthesis of floral flavonols is controlled by differential phylogenetic MYB regulators in *Freesia hybrida*. *New Phytol.* 228, 1864–1879. doi: 10.1111/nph.16818

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Shcherbo, D., Merzlyak, E. M., Chepurnykh, T. V., Fradkov, A. F., Ermakova, G. V., Solovieva, E. A., et al. (2007). Bright far-red fluorescent protein for whole-body imaging. *Nat. Methods* 4, 741–746. doi: 10.1038/nmeth1083

Sneddon, T. P., Li, P., and Edmunds, S. C. (2012). GigaDB: announcing the GigaScience database. *Gigascience* 1:11. doi: 10.1186/2047-217x-1-11

Song, S., Qi, T., Huang, H., Ren, Q., Wu, D., Chang, C., et al. (2011). The Jasmonate-ZIM domain proteins interact with the R2R3-MYB transcription factors *MYB21* and *MYB24* to affect Jasmonate-regulated stamen development in Arabidopsis. *Plant Cell* 23, 1000–1013. doi: 10.1105/tpc.111.083089

Sringarm, K., Potchanasin, P., Sruamsiri, P., and Bangerth, K. F. (2009). Floral induction (FI) in longan (*Dimocarpus longan*, Lour.) trees-The possible participation of endogenous hormones II. Low temperature and potassium chlorate effects on hormone concentrations in and their export out of leaves. *Sci. Hortic.* 122, 295–300. doi: 10.1016/j.scienta.2008.11.031

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Stracke, R., Holtgräwe, D., Schneider, J., Pucker, B., Sörensen, T. R., and Weisshaar, B. (2014). Genome-wide identification and characterisation of R2R3-MYB genes in sugar beet (*Beta vulgaris*). *BMC Plant Biol.* 14:249. doi: 10.1186/s12870-014-0249-8

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

Stracke, R., Jahns, O., Keck, M., Tohge, T., Niehaus, K., Fernie, A. R., et al. (2010). Analysis of production of flavonol glycosides-dependent flavonol glycoside accumulation in *Arabidopsis thaliana plants* reveals *MYB11*-, *MYB12*- and *MYB111*-independent flavonol glycoside accumulation. *New Phytol.* 188, 985–1000. doi: 10.1111/j.1469-8137.2010.03421.x

Stracke, R., Werber, M., and Weisshaar, B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/s1369-5266(00)00199-0

Subramanian, B., Gao, S., Lercher, M. J., Hu, S., and Chen, W. H. (2019). Evolview v3: a webserver for visualization, annotation, and management of phylogenetic trees. *Nucleic Acids Res.* 47, W270–W275. doi: 10.1093/nar/gkz357

Sun, J., Chen, M., Zhu, M., Jiang, Y., Meng, J., Zhao, D., et al. (2018). Cloning, characterization, and expression analysis of three *FAD8* genes encoding a fatty acid desaturase from seeds of *Paeonia ostii*. *Molecules* 23:929. doi: 10.3390/molecules23040929

Sun, W., Ma, Z., Chen, H., and Liu, M. (2019). MYB gene family in potato (*Solanum tuberosum* L.): genome-wide identification of hormone-responsive reveals their potential functions in growth and development. *Int. J. Mol. Sci.* 20:4847. doi: 10.3390/ijms20194847

Teng, S., Keurentjes, J., Bentsink, L., Koornneef, M., and Smeekens, S. (2005). Sucrose-specific induction of anthocyanin biosynthesis in Arabidopsis requires the *MYB75/PAP1* gene. *Plant Physiol.* 139, 1840–1852. doi: 10.1104/pp.105.066688

Tombuloglu, H. (2020). Genome-wide identification and expression analysis of R2R3, 3R- and 4R-MYB transcription factors during lignin biosynthesis in flax (*Linum usitatissimum*). *Genomics* 112, 782–795. doi: 10.1016/j.ygeno.2019.05.017

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., and Leunissen, J. A. (2007). Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res.* 35, W71–W74. doi: 10.1093/nar/gkm306

Van de Peer, Y. (2004). Computational approaches to unveiling ancient genome duplications. *Nat. Rev. Genet.* 5, 752–763. doi: 10.1038/nrg1449

Waheed, S., Peng, Y., and Zeng, L. H. (2020). Identification and characterization of *DlGI* promoter Involved in photoperiod, light Intensity, hormone, and *DlELF4* response from longan. *J. Am. Soc. Hortic. Sci.* 145, 340–348. doi: 10.21273/jashs04946-20

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., and Yu, J. (2010). KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. *Genomics Proteomics Bioinformatics* 8, 77–80. doi: 10.1016/s1672-0229(10)60008-3

Wang, J., Liu, Y., Tang, B., Dai, X., Xie, L., Liu, F., et al. (2020). Genome-wide identification and capsaicinoid biosynthesis-related expression analysis of the R2R3-MYB gene family in *Capsicum annuum* L. *Front. Genet.* 11:598183. doi: 10.3389/fgene.2020.598183

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293

Wang, Y., Zhang, Y., Fan, C., Wei, Y., Meng, J., Li, Z., et al. (2021). Genome-wide analysis of MYB transcription factors and their responses to salt stress in *Casuarina equisetifolia*. *BMC Plant Biol.* 21:328. doi: 10.1186/s12870-021-03083-6

Wilkins, O., Nahal, H., Foong, J., Provart, N. J., and Campbell, M. M. (2009). Expansion and diversification of the Populus R2R3-MYB family of transcription factors. *Plant Physiol.* 149, 981–993. doi: 10.1104/pp.108.132795

Williams, C. E., and Grotewold, E. (1997). Differences between plant and animal Myb domains are fundamental for DNA binding activity, and chimeric Myb domains have novel DNA binding specificities. *J. Biol. Chem.* 272, 563–571. doi: 10.1074/jbc.272.1.563

Winterhagen, P., Hegele, M., Tiyayon, P., and Wunsche, J. N. (2020). Cytokinin accumulation and flowering gene expression are orchestrated for floral meristem development in longan (*Dimocarpus longan* Lour.) after chemical flower induction. *Sci. Hortic.* 270:109467. doi: 10.1016/j.scienta.2020.109467

Winterhagen, P., Tiyayon, P., Samach, A., Hegele, M., and Wunsche, J. N. (2013). Isolation and characterization of *FLOWERING LOCUS T* subforms and *APETALA1* of the subtropical fruit tree *Dimocarpus longan*. *Plant Physiol. Biochem.* 71, 184–190. doi: 10.1016/j.plaphy.2013.07.013

Wong, D. C. J., Schlechter, R., Vannozzi, A., Höll, J., Hmmam, I., Bogs, J., et al. (2016). A systems-oriented analysis of the grapevine R2R3-MYB transcription factor family uncovers new insights into the regulation of stilbene accumulation. *DNA Res.* 23, 451–466. doi: 10.1093/dnares/dsw028

Wu, J., Jiang, Y., Liang, Y., Chen, L., Chen, W., and Cheng, B. (2019). Expression of the maize MYB transcription factor ZmMYB3R enhances drought and salt stress tolerance in transgenic plants. *Plant Physiol. Biochem.* 137, 179–188. doi: 10.1016/j.plaphy.2019.02.010

Yang, Y., Zhang, L., Chen, P., Liang, T., Li, X., and Liu, H. (2020). UV-B photoreceptor UVR8 interacts with *MYB73*/*MYB77* to regulate auxin responses and lateral root development. *Embo J.* 39:e101928. doi: 10.15252/embj.2019101928

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16, 284–287. doi: 10.1089/omi.2011.0118

Zhang, L., Chen, L., and Yu, D. (2018). Transcription factor WRKY75 interacts with DELLA proteins to affect flowering. *Plant Physiol.* 176, 790–803. doi: 10.1104/pp.17.00657

Zhang, X. F., Guo, S., Ho, C. T., and Bai, N. S. (2020). Phytochemical constituents and biological activities of longan (*Dimocarpus longan* Lour.) fruit: a review. *Food Sci. Hum. Wellness* 9, 95–102. doi: 10.1016/j.fshw.2020.03.001

Zhao, F., Zhao, T., Deng, L., Lv, D., Zhang, X., Pan, X., et al. (2017). Visualizing the essential role of complete virion assembly machinery in efficient hepatitis c virus cell-to-cell transmission by a viral infection-activated split-intein-mediated reporter system. *J. Virol.* 91:e1720-16. doi: 10.1128/jvi.01720-16

Zhao, P., Hou, S., Guo, X., Jia, J., Yang, W., Liu, Z., et al. (2019). A MYB-related transcription factor from sheepgrass, *LcMYB2*, promotes seed germination and root growth under drought stress. *BMC Plant Biol.* 19:564. doi: 10.1186/s12870-019-2159-2

Zhao, Y., Xing, L., Wang, X., Hou, Y. J., Gao, J., Wang, P., et al. (2014). The ABA receptor PYL8 promotes lateral root growth by enhancing MYB77-dependent transcription of auxin-responsive genes. *Sci. Signal.* 7:ra53. doi: 10.1126/scisignal.2005051

Zheng, W., Dong, X. M., Zhang, Q. Y., Yu, X. F., Li, W. L., Ji, Y. B., et al. (2020). Identification and characterization of myb genes in *Dimocarpus Longan* Lour. *Bangladesh J. Bot.* 49, 97–104. doi: 10.3329/bjb.v49i1.49100

Zhu, L., Guan, Y., Liu, Y., Zhang, Z., Jaffar, M. A., Song, A., et al. (2020). Regulation of flowering time in chrysanthemum by the R2R3 MYB transcription factor *CmMYB2* is associated with changes in gibberellin metabolism. *Hortic. Res.* 7:96. doi: 10.1038/s41438-020-0317-1

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 May 2022
doi: 10.3389/fpls.2022.852001

Analysis of the β-Glucosidase Family Reveals Genes Involved in the Lignification of Stone Cells in Chinese White Pear (*Pyrus bretschneideri* Rehd.)

*Han Wang[1†], Yingjie Zhang[1†], Xiaofeng Feng[1], Fulei Peng[1], Muhammad Aamir Mazoor[1], Yang Zhang[1], Yu Zhao[1], WenLong Han[1], Jinjin Lu[1], Yunpeng Cao[2] and Yongping Cai[1]**

[1] School of Life Sciences, Anhui Agricultural University, Hefei, China, [2] CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, China

OPEN ACCESS

Edited by:
Yuexue Liu,
Shenyang Agricultural University, China

Reviewed by:
James Robert Ketudat Cairns,
Suranaree University of Technology, Thailand
Ze Peng,
South China Agricultural University, China

***Correspondence:**
Yongping Cai
ypcaiah@163.com

[†]These authors have contributed equally to this work

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

Received: 10 January 2022
Accepted: 31 March 2022
Published: 10 May 2022

Citation:
Wang H, Zhang Y, Feng X, Peng F, Mazoor MA, Zhang Y, Zhao Y, Han W, Lu J, Cao Y and Cai Y (2022) Analysis of the β-Glucosidase Family Reveals Genes Involved in the Lignification of Stone Cells in Chinese White Pear (Pyrus bretschneideri Rehd.). Front. Plant Sci. 13:852001. doi: 10.3389/fpls.2022.852001

BGLU β-glucosidases in glycoside hydrolase family 1 (GH1) are involved in many processes of plant secondary metabolism. In particular, its de-glycosylation function plays an important role in the transport of lignin monolignols. No comprehensive study of the BGLU family in Chinese pear (*Pyrus bretschneideri* Rehd.) has been reported yet. In this study, the 50 BGLU family members from Chinese white pear were identified. Three candidate genes, *PbBGLU1*, *PbBGLU15*, and *PbBGLU16,* that may be involved in lignin synthesis were screened by bioinformatics analysis and qRT-PCR. Subcellular localization showed that all three of these candidate genes were expressed in the extracellular region. Then, we analyzed the functions of *PbBGLU1* and *PbBGLU16. In situ* hybridization analysis showed that *PbBGLU1* transcripts were not only localized to some pulp cell walls, lignin deposition, and stone cell areas of a pear fruit, but that was also a small amount of enrichment in normal pear flesh cells. *PbBGLU16* transcripts were only enriched in lignin deposition and stone cell areas of pear fruit. Enzyme activity analysis revealed that GST-PbBGLU1 and GST-PbBGLU16 had a stronger activity and higher catalytic efficiency for coniferin than syringin. In addition, GST-PbBGLU16 exhibited the higher activity and catalytic efficiency for the two substrates compared with GST-PbBGLU1. The transformation of *PbBGLU1* and *PbBGLU16* into *Arabidopsis* identified that the lignin contents of *Arabidopsis BGLU-45* mutant, *PbBGLU1-RE,* and *PbBGLU16-RE* were not changed than that of wild-type. However, compared with wild-type *Arabidopsis*, the overexpression of the plant's lignin increased in varying degrees. The effect of *PbBGLU16* on the lignin increment was greater than that of *PbBGLU1* in *Arabidopsis*. In pear fruits, with transient overexpression of *PbBGLU1*, the contents of lignin and stone cells were significantly higher ($0.01 < P < 0.05$) than those with empty vector injection pear fruits. After transient expression of *PbBGLU16*, lignin in pear fruit

increased significantly ($0.01 < P < 0.05$) and stone cells showed a very significant difference ($P < 0.01$) compared with the control group. However, RNA interference silenced these two genes in pear fruit, which seemed to have no impression on lignin and stone cells. This study provides a molecular biological basis for improving pear fruit quality at the molecular level.

Keywords: pear, β-glucosidase, gene function, lignin synthesis, stone cells

INTRODUCTION

Pear belongs to the *Rosaceae* family and is one of the most important deciduous fruit tree species worldwide (Wu J. et al., 2012). "Dangshan Su" pear (*Pyrus bretschneideri* cv. Dangshan Su) is a diploid pear variety that originated in Dangshan County, Anhui Province, China. It not only has good flavor and high nutritional value but also has medicinal value and is very popular among people. However, due to its high content of stone cells, its economic value is not high enough (Yan et al., 2014). Stone cells are a kind of specific cell in pear fruit, and they are among the important factors that affect pear fruit processing and fresh consumption. Current research shows that the diameter, number, and density of stone cells have significant effects on pear fruit quality. The greater the diameter and density of the cell mass is, the greater the stone cell content of pear fruit, and the worse the fruit quality is (Zhang et al., 2017). Stone cells are formed by secondary thickening of parenchyma cells, and lignin is one of the crucial components of stone cells (Cheng et al., 2019a; Su et al., 2019).

Lignin is a secondary metabolite. Lignin synthesis, transportation, and deposition are closely related to stone cells development during pear fruit development (Cai et al., 2010; Jin et al., 2013). Structurally, lignin is a phenylpropanoid polymer mainly derived from three kinds of monolignols (*p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol) connected by different chemical bonds (Boerjan et al., 2003; Barros et al., 2015). Glycosylation of these three monolignols is catalyzed by UDP-glucosyltransferase (UGT) into monolignol glucosides (*p*-coumarin, coniferin, syringin) (Cheng et al., 2019b; Perkins et al., 2019), and glycosylation increases the solubility and stability of monolignols, which is beneficial to the transport and storage of monolignols (Gachon et al., 2005). Studies have found that a large number of monolignol glucosides can be stored in the vacuoles of different xylem cells (Steeves et al., 2001; Tsuji and Fukushima, 2004). In-depth research on the transport mechanism of monolignols is scarce, but the academic community currently believes that there are three transport mechanisms for the extracellular secretion of monolignols, namely, passive diffusion (PD), vesicle-related exocytosis, and the use of ABC transport factors or proton coupling to reverse the active transport of ATP (Miao and Liu, 2010; Alejandro et al., 2012; Barros et al., 2015). These monolignol glucosides are then transported to specific parts of the cell wall and hydrolyzed to lignin monolignols under the action of β-glucosidase, ultimately forming guaiacyl (G), syringyl (S), and p-hydroxyphenyl (H) units by laccase and peroxidase [dicotyledons have only guaiacyl (G) and syringyl (S) lignin] (Liu, 2012).

In recent years, an increasing number of species of the β-glucosidase family have been discovered, and the function of the β-glucosidase-encoding gene has become clearer. In 1994, coniferin-specific β-glucosidase was identified in the xylem of two Pinus species (Leinhos et al., 1994; Dharmawardhana et al., 1995). A study in 1999 found that coniferin β-glucosidase catalyzed the hydrolysis of monolignol glucosides to release cinnamyl alcohols for oxidative polymerization to lignin (Dharmawardhana et al., 1999). *At1g61810* (*AtBGLU45*), *At1g61820* (*AtBGLU46*), and *At4g21760* (*AtBGLU47*) in *Arabidopsis* play an important role in the synthesis of lignin (Escamilla-Treviño et al., 2006; Chapelle et al., 2012). *Os4BGlu14*, *Os4BGlu16*, and *Os4BGlu18*, which are members of the rice β-glucosidase family, have been proven to be functional genes involved in rice lignin synthesis (Baiya et al., 2018). In addition, a β-glucosidase was identified from poplar and named "PtrBGLU6"; this gene plays an important role in poplar lignification and cell wall formation (Tsuyama and Takabe, 2015).

In our study, all β-glucosidase family members in pear had been identified and further three β-glucosidase (*PbBGLU1*, *PbBGLU15,* and *PbBGLU16*) candidate genes were considered as functional genes for pear lignin synthesis. We through a series of experiments including *in situ* hybridization, *in vitro* enzyme activity studies, stable expression in *Arabidopsis thaliana* and transient expression in pear fruit. We provided evidence that *PbBGLU1* and *PbBGLU16* can act in the formation of lignin and stone cell development in pear fruit.

MATERIALS AND METHODS

Identification of *BGLU* Genes in Five *Rosaceae* Species

The Chinese white pear genome database was obtained from the Pear Genome Project.[1] To identify putative BGLU genes from Chinese white pear, several approaches were employed. We downloaded the HMM profile (PF00232) from the Pfam website[2] (Finn et al., 2010) and searched for candidate genes using the HMM with *E*-values < 1*e*-10 from the Chinese white pear genome (Zhou et al., 2016). After performing multiple sequence alignment, the retrieved domain sequences of the pear

Abbreviations: BGLU, β-glucosidase; GH1, glycoside hydrolase family 1; HMM, hidden Markov model; NJ, neighbor-joining; qRT-PCR, real-time PCR; ABC transporter, ATP-binding cassette transporter.

[1]http://gigadb.org/dataset/10008

[2]https://pfam.xfam.org/

are aligned, and the HMM of the BGLU genes family of the pear is constructed again. Using the constructed new hidden Markov model with *E*-values < 1*e*-10 to blasted against the Chinese pear genome for carryout PbBGLU genes. The online site SMART[3] (Letunic et al., 2012) was used to determine the domain PF00232 of BGLU. The online analysis tool EXPASY[4] was used to predict and analyze the physical and chemical properties of the obtained BGLU amino acids sequences.

Phylogenetic Conserved Motif and Exon-Intron Structures Analyses of BGLU Genes in Pear

Sequence alignment of *PbBGLUs* proteins was done by using the MUSCLE method in MEGA 7.0 software. The phylogenetic tree was constructed with MEGA 7.0 software using the neighbor-joining (NJ) (bootstrap = 2,500) (Kumar et al., 2016). The conservative motifs were analyzed by MEME[5] software the maximum value of the motif was set to 20, and the motif length was set between 6 and 50 (Bailey et al., 2015). The Gene Structure Server[6] was used to generate exon-intron maps (Hu et al., 2015).

Phylogenetic analysis of *BGLU* gene family in *Arabidopsis*, *P. bretschneideri, Oryza sative,* and *Populus*. The software MEGA 7.0 was used to construct the phylogenetic tree. The amino-acid sequences of *Arabidopsis*, *O. sative,* and *Populus* were obtained from phytozome database.[7]

Plant Materials, Treatments, and Growth Conditions

The materials were from "Dangshan Su" pear (*P. bretschneideri* cv. Dangshan Su) with the same growing environment for 40 years in Dangshan County, Anhui Province, China. We collected the mature leaves, buds, stem segments, and flowers on March 30, 2021; and collected the fruit on April 15 (15 days after flowering), 2021; May 14 (39 DAF), 2021; May 22 (47 DAF), 2021; June 4 (55 DAF), 2021; June 12 (63 DAF), 2021; June 28 (79 DAF), 2021; and August 30 (145 DAF), 2021. Each sample was frozen in liquid nitrogen and brought back to the laboratory for use. All samples were stored at–80°C for subsequent RNA extraction.

Arabidopsis thaliana L. wild type was purchased from the American Arabidopsis Biological Resources Center. *Arabidopsis At1g61810 bglu45-2* mutant seeds corresponding to the Salk_117269 flanking sequence tag were ordered from the Salk Institute Genomic Analysis Laboratory. Homozygous *BGLU45-2* mutant plants were isolated by genotype analysis with *Arabidopsis* BGLU45 gene and T-DNA specific primers, as described by Chapelle et al. (2012). The *Arabidopsis* plants were cultivated in a growth chamber at 23°C under 16/8 h light/dark cycles (constant illumination 100 μEm-2s-1). Ten replicates of each line were planted, and three similarly growing plants were collected for further analysis.

[3]http://smart.embl.de/

[4]http://web.expasy.org/compute_pi/

[5]https://meme-suite.org/meme/doc/meme.html

[6]http://gsds.cbi.pku.edu.cn

[7]https://phytozome.jgi.doe.gov/pz/portal.html#

RNA Extraction, Reverse Transcription PCR, and Real-Time PCR Analysis

The total RNA of all the used "Dangshan Su" pear fruits in this study was extracted using a Plant RNA extraction kit from Tiangen (Beijing, China). Total RNA (1 μg) from each sample was used in reverse transcription. First-strand cDNA was synthesized with a PrimeScript^{TM} RT reagent kit with a gDNA Eraser kit (TaKaRa, Kyoto, Japan). All RT-qPCR primers were designed using Primer Premier 5.0 software (**Supplementary Table 3**). We chose the tubulin (No. AB239680.1) served as the reference gene for RT-qPCR analysis (Wu T. et al., 2012). The RT-qPCR was performed with a CFX96 Touch^{TM} Real-Time PCR Detection System (BIO-RAD), each sample was subjected to three biological replicates. The relative changes in gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Subcellular Localization of *PbBGLU1, PbBGLU15,* and *PbBGLU16*

The three genes RT-PCR primers were designed using Primer Premier 5.0 software (**Supplementary Table 4**). The full-length CDS of *PbBGLU1*, *PbBGLU15,* and *PbBGLU16* without stop codon were cloned based on genomic information and constructed into pCAMBIA1305 vectors (Clontech, Beijing, country-region China) between the *Xba*I and *Bam*HI (NEB) sites which have both CaMV35S promoter and GFP gene. After electroporation of these constructions into *Agrobacterium tumefaciens EHA105*, using pCAMBIA1305 vector as a negative control. The infection solution was injected into the epidermis of *Nicotiana tabacum* leaves, after culturing in the dark for 3 days, the glass slide was made and placed under the FV1200 laser confocal microscope (Olympus Corporation, Tokyo, Japan) to observe the distribution of fusion protein.

In situ Hybridization

Segments of the 39 DAF fresh pear fruit were fixed in *in situ* hybridization fixed solution overnight. Pear pulp segments fixed: Take out the segment, wash, clean, put in the fixed fluid (DEPC) immediately to fix above 12 h. Dehydration: The segment is dehydrated by gradient alcohol, paraffin, embedding, and vacuum pumping in the dehydration process. Section: The paraffin is sliced through the slicer, the piece of the slicing machine, and the 62–degree oven roast for 2 h. Dewaxing and dehydration: Soak sections in 2 changes of xylene, 15 min each. Dehydrate in 2 changes of pure ethanol for 5 min each. Then, followed respectively by dehydrating in gradient ethanol of 85 and 75% ethanol 5 min each. Wash in DEPC dilution. Digestion: Mark the objective tissue with a liquid blocker pen, according to the characteristics of tissues, add proteinase K (20 ug/ml) working solution to cover objectives and incubate at 37°C for 22 min. Washing with pure water, then wash three times with PBS (pH 7.4) in a rocker device, 5 min each. Hybridization: Discard the pre-hybridization solution, add the probe hybridization solution, concentration, and incubate the section in a humidity chamber and hybridize overnight. Add the Alkaline Phosphatase-conjugated IgG Fraction Monoclonal

Mouse Anti-Digoxin Antibody: (anti-DIG-AP): Remove the blocking solution and add anti-DIG-HRP. Incubate at 37°C for 40 min, then wash sections in TBS four times for 5 min each. NBT/BCIP developing: dry sections slightly, add fresh prepared NBT/BCIP (Thermo Scientific 1-Step NBT/BCIP) chromogenic reagent to marked tissue. Manage reaction time by observing in microscopy until positive expression appears shows blue-purple. Then stop developing reaction by a wash in running tap water. The probe sequence was unique to the PbBGLU1 and PbBGLU16 locus (**Supplementary Table 4**) and resulted in a single hit when used as a quarry to BLAST the Chinese white pear genome. *In situ* hybridization probes were synthesized by Sangon Bio Shang hai and labeled with Digoxigenin.

Construction of PbBGLUs Expression Vector and Induced Purification of Recombinant Protein

The RT-PCR primers were designed using Primer Premier 5.0 software (**Supplementary Table 4**). The full-length coding sequence of *PbBGLU1* and *PbBGLU16* were cloned and then expressed in pGEX4T-1 vector (GE Healthcare, Chicago, IL, United States) between the SamI and *Not*I (NEB) sites. Two recombinant pGEX4T-1-PbBGLUs were transformed into *Escherichia coli* BL21 (DE3) Competent Cells. Add *Escherichia coli* BL21 (DE3) cells carrying pGEX4T-1-PbBGLU1 and pGEX4T-1-PbBGLU16 plasmid to 100 ml of LB liquid containing ampicillin for shaking. Expand the culture to an *OD* value of about 0.6, add isopropyl-β-D-thioacetamide IPTG to a final concentration of 1 mM/L to induce the fusion expression of the recombinant protein, culture with shaking at 16°C for 24 h, and collect the cells by centrifugation at 4°C. The centrifuged cells were suspended in PBS (pH 7.2–7.4) buffer, using an ultrasonic cell disruptor for disruption, then centrifuge to take the supernatant and discard the pellet. Using affinity chromatography resin with GST tag GST· Bind^{TM} resin (Novagen) to Purify protein. The Elution Buffer (pH 8.0) for GST-Sefinose (TM) Resin (containing reducing glutathione reagent at a concentration of 5 mM) Eluent was used for elution.

Recombinant Protein Enzyme Assays

To analyze the activity of recombinant protein GST-PbBGLU1 and GST-PbBGLU16, 10 μg of purified protein was incubated at 35°C with 20 μL buffer (50 mM MgSO4, 200 mM KCL, 100 mM PBS pH 7.2–7.4) and 1 mM substrates (coniferin and syringin). The water was added to a final volume of 50 μL. After 1 h of reaction, add 50 μL methanol termination reaction. All reactions were supplemented with 0.1% (V/V) β- Mercaptoethanol. Then the reaction results were analyzed by HPLC.

To calculate the K_m, V_{max}, and k_{cat} value of the recombinant protein, under the same original reaction conditions, the substrates (coniferin and syringin) were set to 10 concentrations (0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, and 0.6 mM). Then the product concentration C is calculated by an external standard method, and then the reaction rate V is further calculated. Finally, use the double reciprocal method to plot to calculate K_m and V_{max}.

Genetic Transformation of *Arabidopsis* and Transient Transformation of Pear Fruit

The RT-PCR primers were designed using Primer Premier 5.0 software (**Supplementary Table 4**). The full-length coding sequence of *PbBGLU1* and *PbBGLU16* were cloned from "Dangshan Su" pear first-strand complementary DNA (cDNA) with PrimeSTAR® Max DNA Polymerase (Takara, Japan). The fragment was cloned into pCAMBIA1301-35S binary vectors between the *Bam*HI and *Xba*I (NEB) sites and used in plant transformation. Wild type and *Arabidopsis At1g61810 BGLU45-2* mutant were used for transformation with *A. tumefaciens* GV3101 carrying the above binary plasmid using the floral dip method.

The full-length coding fragment was cloned into pCAMBIA1301-35S binary vectors and used in the transient injection of pear fruit overexpression. Hairpin constructs were based on the pRNAiDE001 vector (Heneghan et al., 2007; Eastwood et al., 2008) where self-complementary sense (PbBGLU1: 905-1220bP; PbBGLU16: 832-1183bP) and antisense sequences are separated by a non-functional sequence (loop), then construct this fragment into the pCAMBIA1301 vector with the 35S promoter (**Supplementary Figures 1**, **2**). Pear trees (40-year-old) and 39 days after flowering young fruits were selected for transient injection. The constructed pCAMBIA1301-35s binary vectors of *PbBGLUs* overexpression and *PbBGLUs* RNA interference gene were injected on the left of the fruit, and the empty pCAMBIA1301-35s binary vectors were injected on the right for control. The method followed a previously reported protocol (Zhang et al., 2021) nine fruits were injected with each construct in an experiment that was repeated three times independently. After 10 days, the transient injected fruits were picked up.

Lignin Determination and Histochemical Staining of *Arabidopsis* and Pear Fruits

We collected the *Arabidopsis* plants for approximately 6 weeks, removed the leaves, and dried them in an oven at 65°C for 48 h. The lignin content of the *Arabidopsis* plants was estimated following the method of Anderson et al. (Anderson et al., 2015).

We obtained the inflorescence stem segments (young stem regions) of the transgenic and wild-type *Arabidopsis* plants for approximately 6 weeks. The samples were placed in a solution containing 95% methanol, 70% (v/v) ethanol and glacial acetic acid for 12 h and embedded in paraffin for sectioning with the pathology slicer (RM 2018). Plant tissue sections were stained following standard phloroglucinol staining protocols (Pradhan and Loqué, 2014).

After the transiently injected pear fruits was brought back to the laboratory, samples 5 g pulp from 2.0 mm under the peel to 0.5 mm outside the core, was collected and frozen at –80°C for 24 h, homogenized with a high-speed homogenizer for 3 min, rotating speed was 20,000 RPM/min, added water and stood for a while, then poured out the upper suspension, repeated several times until the upper layer was clear, then filtered and dried stone cells were weighed, repeated for 3 times, stone cell

content = measured stone cell dry weight/5 × 100%. The lignin content was measured using the Klason method (Raiskila et al., 2007). A small amount (0.2 g) of stone cells was extracted with 15 ml of 72% H_2SO4 at 30°C for 1 h, combined with 115 ml of distilled water, and boiled for 1 h. The volume was kept constant during boiling. The liquid mixture was filtered and the residue was rinsed with 500 ml of hot water, air-dried and weighed. The lignin content was shown as a percentage (calculated lignin content/dry weight of stone cells × 100%) (Zhang et al., 2021).

Targeted Metabolite Determination

HPLC (Thermo Scientific UltiMate 3000 HPLC) analysis was carried out using a Columbus (Thermo Scientific 5-μm C18 column; 250×4.60 mm). Acetonitrile (solvent A) and H_2O (solvent B) with gradient of 10–30% acetonitrile in water (all solutions contained 0.1% trifluoroacetic acid) with a flow rate at 1 ml min^{-1} over 35 min was used. Each peak on the chromatogram was scanned between 200 and 400 nm (photodiode array profile) and was integrated at 264, 280, and 324 nm. The data were acquired and analyzed using the software ChromQuest version. Extraction of tissue was carried out as described in Lanot et al. (2006).

RESULTS

Identification of *BGLU* Genes in *Pyrus bretschneideri*

The BGLU family conserved domain (Pfam: PF00232) was used as a target query sequence, and the corresponding hidden Markov model was obtained from the Pfam website (see text footnote 2) which was used to identify the *BGLU* members. SMART (see text footnote 3) was used to check whether there are characteristic domains, with the redundant and repeated sequences ultimately removed. As a result, we identified 50 *BGLU* genes from Chinese white pear (*P. bretschneideri*). The detailed information on the gene ID, gene name, chromosomal location, protein structure, and characteristics of the corresponding BGLU proteins are listed in **Supplementary Table 1**.

Phylogenetic Analyses of the Conserved Motif and the Exon–Intron Structure of *BGLU* Genes in Chinese White Pear

Inclusive of all *PbBGLU* genes, a phylogenetic tree was constructed with MEGA 7.0 software using the neighbor-joining (NJ) method (bootstraps = 1,000), which divided them into 8 groups (I, II, III, IV, V-A, V-B, V-C, V-D) with supported bootstrap values (**Figure 1**). We used MEME online software to analyze the amino acid sequence of *PbBGLU* genes and identified 20 motifs among all the sequences (**Supplementary Table 2** and **Figure 1**). All the *PbBGLU*s contain motif1 (CTGGBSATEPYJVAHHQLLAHAAAVKLYREKYQA), and only two genes do not contain motif9 (WFEPASES KEDKAAALRALDF), suggesting that motif1 and motif9 are conserved gene motifs of *PbBGLU*s family. Only group I, group II, and group III genes contain motif19, whereas the other groups do not, and no members of group V-D contain motif2, it is speculated that these motifs may have special significance. Previous studies implied that gene structural diversity can lead to the evolution of multi-gene families (Dong et al., 2019). To better characterize and understand the structural diversity of the *PbBGLU*s, gene exon-intron analysis was carried out (**Figure 1**). Notably, *PbBGLU23* and *PbBGLU24* have only one exon, whereas the others contain 7 or more exons; *PbBGLU34* contains the most exons (21), and most genes had between 10 and 15 exons. These results indicate the possible occurrence of alternative splicing during gene evolution, which in turn leads to functional differences between these closely related *PbBGLU*s (Cao et al., 2017; Dong et al., 2019).

To further understand the evolutionary relationship and related functions of *PbBGLU*s, a phylogenetic tree comprising *BGLU*s from *A. thaliana* (Xu et al., 2004), *O. sative* (Opassiri et al., 2006), *poplar* (Tsuyama and Takabe, 2015), and *P. bretschneideri* was constructed (**Figure 2**). We marked the position of the eight groups in pear on this phylogenetic tree and found that group I, group II, and group IV contained only *PbBGLU*s but not *AtBGLU*s, *Osbglu*s, and *PtrBGLU*s, indicating that some changes occurred among the *BGLU*s of different species during the evolutionary process (Cao et al., 2016a,b). *Arabidopsis AtBGLU17* (Roepke et al., 2017) was related to the use of *Arabidopsis* flavonoids which was clustered together with *PbBGLU19* (group III), and it is speculated that *PbBGLU19* might be involved in the hydrolysis of flavonoid glycosides in pear. *PbBGLU* group V-D clustered together with *AtBGLU42*, and which encodes an MYB72-dependent key regulator of rhizobacterium-induced systemic resistance and modulates iron deficiency responses in *Arabidopsis* roots (Zamioudis et al., 2014). *AtBGLU1-6*, which plays a role in the accumulation of flavonoids in *Arabidopsis* (Ishihara et al., 2016), clustered together with *PbBGLU* group V-A genes, speculating that *PbBGLU29* and *PbBGLU30* may have similar functions in pear. The functions of *AtBGLU45* and *AtBGLU46* in *Arabidopsis* (Escamilla-Treviño et al., 2006; Chapelle et al., 2012); *Os4BGlu14*, *Os4BGlu16*, and *Os4BGlu18* in *O. sative* (Baiya et al., 2014; Baiya et al., 2018); *PtrBGLU6* in *poplar* (Tsuyama and Takabe, 2015) have been clarified, and their products catalyze the hydrolysis of monolignol glucosides during lignification. The three genes (*PbBGLU1*, *PbBGLU15,* and *PbBGLU16*) in *PbBGLU* group V-C clustered together with these functional genes, speculating that *PbBGLU1*, *PbBGLU15,* and *PbBGLU16* may be able to affect lignification by catalyzing the hydrolysis of monolignol glucosides in pear fruit, which in turn affects the development of pear stone cells.

Expression Characteristics of Chinese White Pear BGLU Genes

We constructed a phylogenetic tree comprising *BGLU*s genes from *A. thaliana* (Xu et al., 2004), *O. sative* (Opassiri et al., 2006), *poplar* (Tsuyama and Takabe, 2015), and *P. bretschneideri*, and found that *PbBGLU1*, *PbBGLU15,* and *PbBGLU16* in the group V-C *PbBGLU*s of pear clustered together with *AtBGLU45*, *AtBGLU46, Os4BGlu14*, *Os4BGlu16*, *Os4BGlu18,* and *PtrBGLU6* functional genes, which can hydrolyze monolignol glucosides

during lignification (Escamilla-Treviño et al., 2006; Chapelle et al., 2012; Baiya et al., 2014; Tsuyama and Takabe, 2015; Baiya et al., 2018). To verify whether *PbBGLU1*, *PbBGLU15,* and *PbBGLU16* have similar functions, we selected all *PbBGLU* genes in group V (a total of 11 *PbBGLU*s) for qRT-PCR analysis. The expression profiles of these genes in different tissues of pear (leaves, buds, stems, flowers) and pear fruit at different developmental stages (15, 39, 47, 55, 63, 79, and 145 days after flowering) were also investigated (**Figure 3**). It found that *PbBGLU43* was not expressed in different tissue parts or at different developmental stages of the fruit. The expression levels of the other ten *PbBGLU*s were generally high in leaves, buds, flowers, and young fruits (15, 39, 47, and 55 days after flowering). This result was consistent with those of previous research showing that β-glucosidase can be found mainly in young vegetative parts (Kristoffersen et al., 2000; Biely et al., 2003). The expression levels of the *PbBGLU1*, *PbBGLU15,* and *PbBGLU16* genes were relatively high in the leaves and flowers as compared to buds and stems. All three genes were expressed at different stages of fruit development, and their expression showed a trend of first increasing then decreasing, with a similar trend of the content for lignin and stone cells in "Dangshan Su" pear fruit (Cai et al., 2010; Jin et al., 2013). The expression patterns of these three genes were also similar to those of the key genes involved in the regulation of the lignin synthesis pathway (Xie et al., 2013). These results suggested that the three genes (*PbBGLU1*, *PbBGLU15*, and *PbBGLU16*) in the group V-C *PbBGLU*s might be functional genes involved in the hydrolysis of monolignol glucosides in the pear lignin synthesis pathway.

Subcellular Localization of PbBGLU1, PbBGLU15, and PbBGLU16

As early as 1995, it found that *Pinus contorta* β-glucosidase activity occurred in the cell wall of secondary xylem tissue (Dharmawardhana et al., 1995). *AtBGLU45* and *AtBGLU46* of *A. thaliana* were also expressed in the cell wall (Chapelle et al., 2012). The rice Os4BGlu16 tagged with a C-terminal GFP was transiently expressed in tobacco leaf epithelial cells, which revealed that the GFP signal was exclusively localized

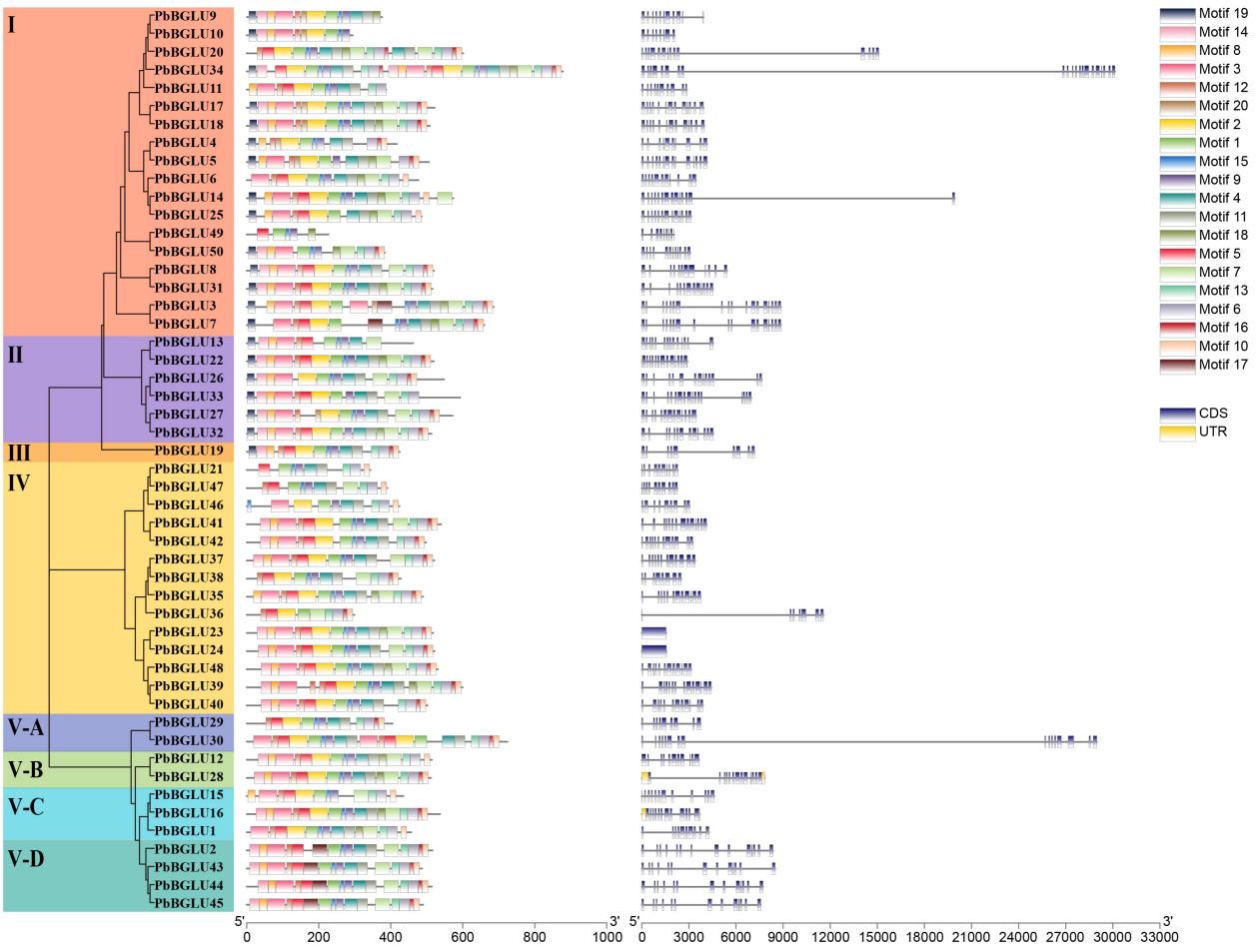


FIGURE 1 | Phylogeny, conserved motifs and exon-intron structure of *BGLU* genes in pear. MEME tools were used to identify motifs. Different colors represent different motifs. GSDS online website was used to generate structure. The legend is located at the bottom right of the figure. Scale represents the length of the DNA sequence.

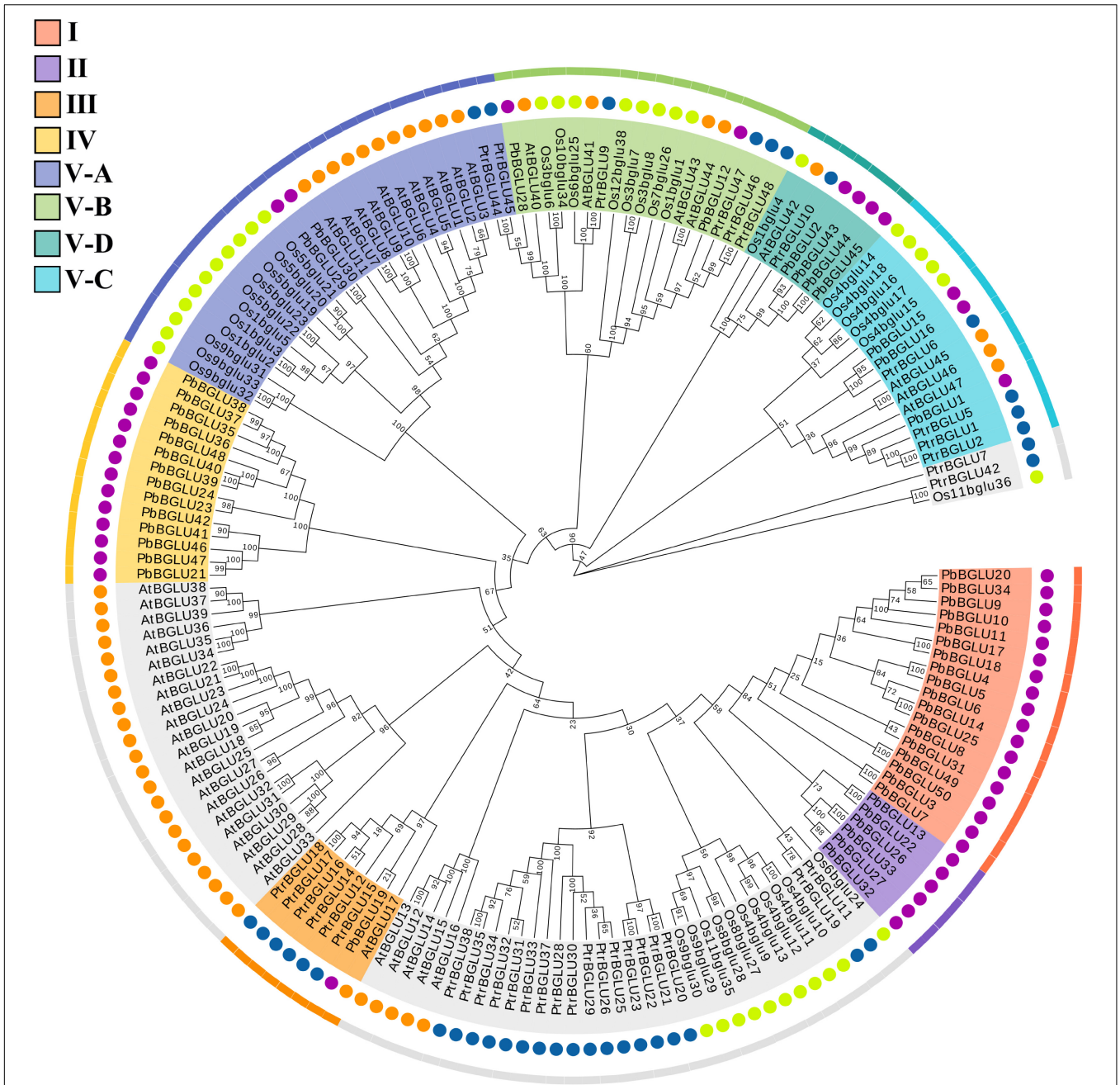


FIGURE 2 | Phylogenetic analysis of *BGLU*s from *Pyrus bretschneideri*, *Arabidopsis thaliana*, *Oryza sative,* and *poplar*. The neighbor-joining (NJ) method was used to construct the phylogenetic tree. Pb: *Pyrus bretschneideri*; Os: *Oryza sative*; Ptr: *Poplar* L.; At: *Arabidopsis thaliana*.

to the plasma membrane and extracellular space in both intact and plasmolyzed cells (Baiya et al., 2018). It is consistent with apoplastic or cell wall localization. In this study, to explore the subcellular localization of *PbBGLU1*, *PbBGLU15,* and *PbBGLU16*, PbBGLU-GFP expression vectors were constructed and transformed into *N. tabacum*. As shown in **Figure 4**, green fluorescence signals from the expressed PbBGLU1-GFP, PbBGLU15-GFP, and PbBGLU16-GFP fusion constructs were specifically distributed in the extracellular region.

In situ Hybridization Analysis of *PbBGLU1* and *PbBGLU16* in 39 DAF Pear Fruit

To confirm the *PbBGLU1* and *PbBGLU16* gene expression associated with those stone cells undergoing lignin deposition and secondary wall thickening, we carried out RNA *in situ* hybridization (Wu and Wagner, 2012) on 39 DAF (days after flowering) pear fruit. The results of *PbBGLU1* transcripts *in situ*

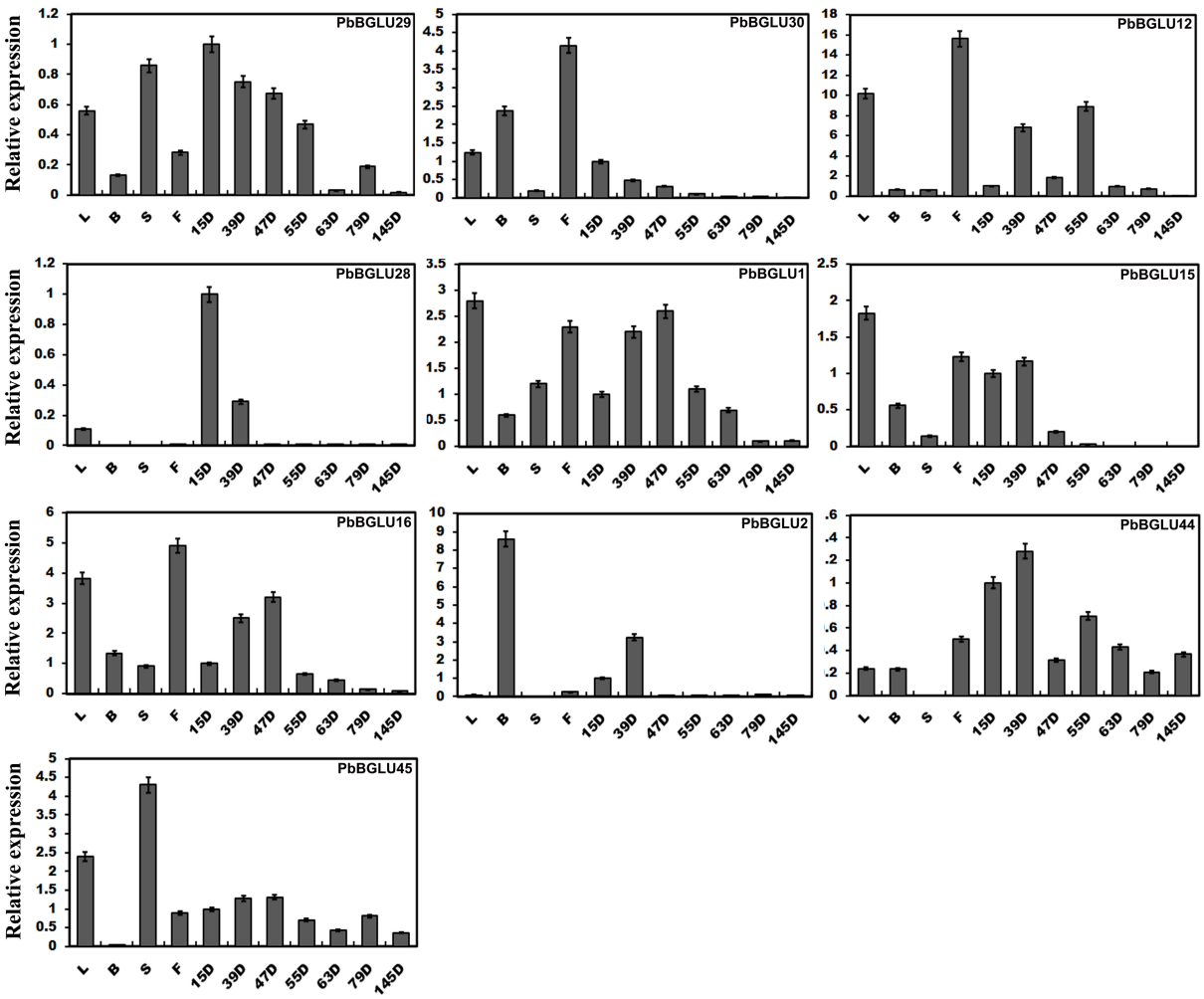


FIGURE 3 | Expression levels of PbBGLU genes in different plant tissues. 15D, 39D, 47D, 55D, 63D, 79D, and 145D correspond to seven different developmental stages (days after flowering) of pear fruit. In addition, leaf, bud, stem, and flower were represented by L, B, S, and F, respectively. The value on the left Y-axis indicates the relative gene expression levels.

hybridization with *PbBGLU1* anti-sense probe showed that *PbBGLU1* transcripts were not only localized to some pulp cell walls, lignin deposition and stone cell areas of a pear fruit, there was also a small amount of enrichment in normal pear flesh cells (**Figure 5A**). This result shows that *PbBGLU1* is not only involved in lignin biosynthesis but also may be involved in other metabolic activities during pear fruit development. However, *PbBGLU16* transcripts were only enriched in lignin deposition and stone cell areas of pear fruit (**Figure 5C**), indicating that it may be a specific gene for lignin synthesis in pear fruit. Hybridization with control *PbBGLU1* and *PbBGLU16* sense probe showed no labeling as expected (**Figures 5B,D**).

Enzymatic Assays of Recombinant PbBGLU1 and PbBGLU16 *in vitro*

PbBGLU1 and *PbBGLU16* were cloned into pGEX4T-1 vector, and then the constructed vector was induced and expressed in BL21 (ED3) *E. coli*. A target protein GST-PbBGLU1 with a weight of 78.07 kDa and another target protein GST-PbBGLU16 with a weight of 87.23 kDa obtained (GST: 26kDa; *PbBGLU1*: 52.07 kDa; *PbBGLU16*: 61.23 kDa). The sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to analysis and verification (**Supplementary Figure 3**). To verify whether the recombinant protein has the ability to glycosylate the monolignol glucosides coniferin, syringin. We added recombinant protein, coniferin, and syringin and buffer in the certain reaction system. After the reaction for a while, methanol was added to terminate the reaction and then put into the HPLC for detection. The standard of coniferyl alcohol, sinapyl alcohol, coniferin and syringin were put into the HPLC for reference (**Figures 6A,B**). The results showed that after the recombinant protein was added, two peaks of the substrate (coniferin and syringin) and product (coniferyl alcohol and sinapyl alcohol) were detected by HPLC. This phenomenon indicates that the recombinant protein of GST-PbBGLU1 and GST-PbBGLU16 were active against both substrates (coniferin and syringin).

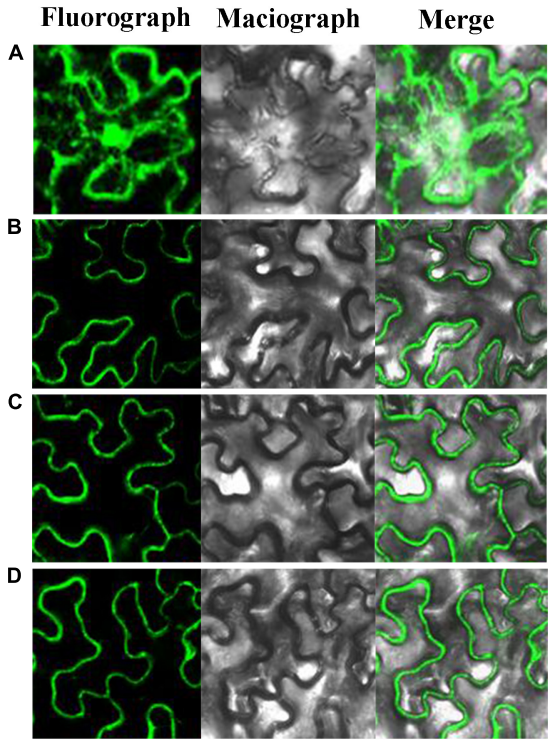


FIGURE 4 | Subcellular localization of PbBGLU1-GFP, PbBGLU15-GFP, and PbBGLU16-GFP fusion protein. **(A)** GFP control; **(B)** Subcellular localization of PbBGLU1-GFP fusion protein; **(C)** Subcellular localization of PbBGLU15-GFP fusion protein; **(D)** Subcellular localization of PbBGLU16-GFP fusion protein.

To calculate the kinetic parameters of the recombinant protein for coniferin and syringin enzyme, 10 substrates with different concentrations were set in this study. The K_m values, V_{max} and k_{cat} of GST-PbBGLU1 and GST-PbBGLU16 were calculated through the double reciprocal method. The results are shown in **Supplementary Figure 4** and **Table 1**. Compared with syringin, GST-PbBGLU1 and GST-PbBGLU16 had a stronger activity and faster catalytic efficiency for coniferin. In addition, compared with GST-PbBGLU1, GST-PbBGLU16 shows higher catalytic activity and catalytic efficiency for the two substrates.

Analysis of Transgenic *Arabidopsis* With *PbBGLU1* and *PbBGLU16* Gene

The *PbBGLU1* and *PbBGLU16* were transferred into *Arabidopsis BGLU-45* mutants (*PbBGLU1-RE* and *PbBGLU16-RE*) and wild-type *Arabidopsis* (*PbBGLU1-OE* and *PbBGLU16-OE*), reverse transcription PCR validation of *Arabidopsis* is shown in **Figures 7A,B**. To directly observe the *in situ* distribution of lignin in inflorescence stems of these *Arabidopsis* plants, we performed the Wiesner (phloroglucinol-HCl) histochemical staining (Pradhan and Loqué, 2014) of inflorescence stems of *Arabidopsis* plants. The Wiesner staining results indicated that the xylem and intravascular fibers in the stems of *Arabidopsis BGLU-45* mutants showed no increase to any extent compared with wild-type *Arabidopsis* (**Figures 7C-2,D-2**), this result is consistent with the results of previous studies (Chapelle et al., 2012). Then, we observed the stained sections of *PbBGLU1-OE* and *PbBGLU16-OE Arabidopsis* and the Wiesner staining results indicated that the xylem and intravascular fibers in the stems of *Arabidopsis PbBGLU1-OE* and *PbBGLU16-OE* showed stronger phloroglucinol staining compared to the wild-type plants (**Figures 7C-4,D-4**). According to the staining results, we can also find that *PbBGLU16-OE* xylem and intravascular fibers *Arabidopsis* increase more strongly than *PbBGLU1-OE Arabidopsis*. Subsequently, we measured the lignin content of wild-type, *Arabidopsis BGLU-45* mutant, *PbBGLU1-RE,* and *PbBGLU16-RE*, *PbBGLU1-OE,* and *PbBGLU16-OE Arabidopsis* by acetyl bromide method (Anderson et al., 2015; **Figures 7E,F**). It found that the lignin content of *Arabidopsis BGLU-45* mutant, *PbBGLU1-RE* and *PbBGLU16-RE* was no change than that of wild-type. However, compared with wild-type *Arabidopsis* the overexpression plants lignin increased in varying degrees, the effect of *PbBGLU16* on the increase of lignin in *Arabidopsis* is greater than that of *PbBGLU1*.

Targeted Metabolite Determination

In order to further explore how *BGLU-45* mutants and overexpression of *PbBGLU1* and *PbBGLU16* affect the lignin of *A. thaliana*, we measured the *A. thaliana* contents of coniferin and syringin by HPLC. The result showed that coniferin and syringin increased significantly in *BGLU-45* mutant compared with the wild type (**Table 2**), but the lignin did not increase. This phenomenon occurred probably because coniferin and syringin are not precursors of lignin synthesis and usually exist in vacuoles as transport or storage forms. This is also probably the reason why there is no significant change in *Arabidopsis* lignin after knocking out *BGLU45* in previous research (Chapelle et al., 2012; Baiya et al., 2018). Then, the contents of coniferin and syringin rising phenomenon have been restored through transferred *PbBGLU1* and *PbBGLU16* into *Arabidopsis BGLU-45* mutant (**Table 2**). However, the lignin monolignol glycosides of transgenic *Arabidopsis* after overexpression of *PbBGLU1* and *PbBGLU16* decreased to some extent, but the decrease was not obvious, it may be because the decrease of lignin monolignol glycosides will increase the expression of *UGT*.

Transient Transformation of *PbBGLU1* and *PbBGLU16*, *PbBGLU1-RNAi* and *PbBGLU16-RNAi* in Pear Fruit

Since the stable genetic transformation system of pears is in its infancy, it is difficult to perform a functional analysis of the *PbBGLU1* and *PbBGLU16* in pears. Therefore, we chose the method of transient expression by injection for analysis. We injected transiently the *PbBGLU1*, *PbBGLU16,* and *PbBGLU1* and *PbBGLU16* RNA interference in the 39 DAF (days after flowering) fruit of the pear tree. By Wiesner (phloroglucinol-HCl) histochemical staining (Pradhan and Loqué, 2014), the results show (**Figures 8B,D**) that when *PbBGLU1* and *PbBGLU16* were silenced by RNA interference, the change of pear lignin and stone cells was not obvious compared with the empty vector injection. But after *PbBGLU1* and *PbBGLU16* gene were overexpressed in

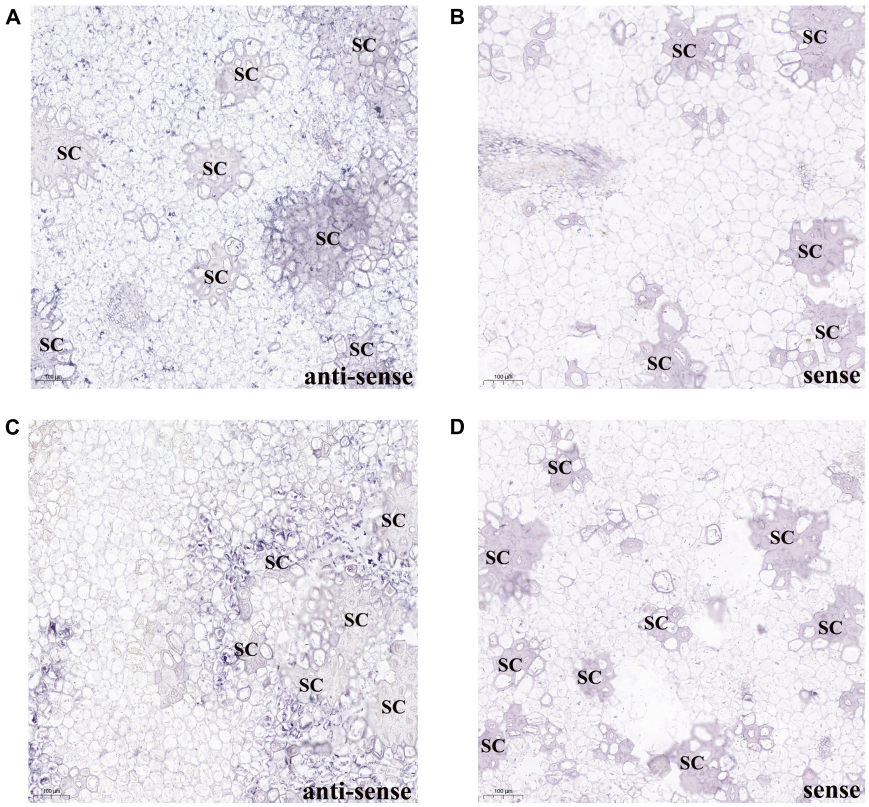


FIGURE 5 | RNA *in situ* hybridization illustrating PbBGLU1 and PbBGLU16 transcript localization in 39 DAF (day after flower) pear fruit. **(A,B)** PbBGLU1 RNA *in situ* hybridization in 39 DAF (day after flower) pear fruit. **(C,D)** PbBGLU16 RNA *in situ* hybridization in 39 DAF (day after flower) pear fruit; Sections were taken through the flesh and probed with anti-sense **(A,C)** and sense probes **(B,D)** and imaged at 10 magnification. The positive expression of BCIP/NBT was blue-purple; the sense probes is for negative control; the bar = 100 μm; SC, stone cells.

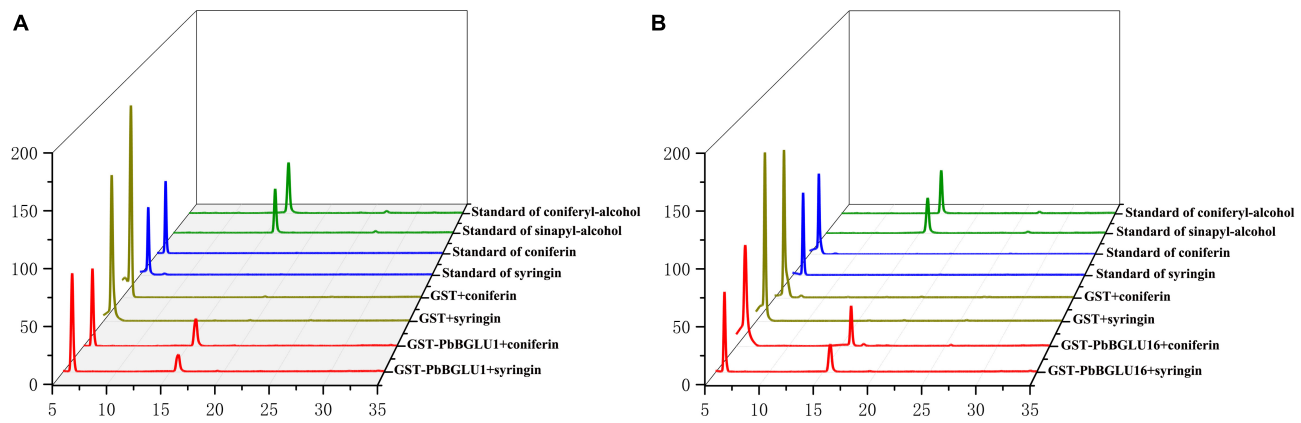


FIGURE 6 | Enzyme activity of GST-PbBGLU1 and GST-PbBGLU16 recombinant protein toward coniferin and syringin. **(A)** HPLC analysis of standards and reaction products of GST-PbBGLU1 recombinant protein. **(B)** HPLC analysis of standards and reaction products of GST-PbBGLU16 recombinant protein. Reaction conditions: 10 μg of purified protein was incubated at 35°C with 20 μL buffer (50 mM MgSO4, 200 mM KCL, 100 mM PBS pH 7.2–7.4) and 1 mM substrates (coniferin and syringin), the water was added to a final volume of 50 μL. After 1 h of reaction, add 50 μL methanol termination reaction.

the 39 DAF pear fruit, it found that the lignin and stone cells increased compared with the empty vector injection control in the pear fruit (**Figures 8A,C**). Moreover, compared with the transient expression of *PbBGLU1*, *PbBGLU16* had a stronger effect on the lignin and expression cells of pear fruit. To explore the specific effects of overexpression and RNAi of *PbBGLU1* and *PbBGLU16* on stone cells and lignin of pear fruit, we measured the contents of stone cells and lignin in pear fruit

TABLE 1 | Enzyme kinetics of PbBGLU1 and PbBGLU16.

Protein	Substrates	K_m (μM)	V_{max} (μM/min·mg)	k_{cat} (S^{-1})	k_{cat}/K_m (mM^{-1} S^{-1})
pGEX4T-1-PbBGLU1	coniferin	102.59 ± 3.89	1.76 ± 0.04	0.23 ± 0.01	2.24 ± 0.02
	syringin	100.77 ± 6.51	1.57 ± 0.05	0.20 ± 0.02	1.98 ± 0.07
pGEX4T-1-PbBGLU16	coniferin	98.01 ± 4.98	3.73 ± 0.10	0.54 ± 0.01	5.52 ± 0.17
	syringin	119.48 ± 8.17	2.46 ± 0.09	0.36 ± 0.01	3.02 ± 0.12

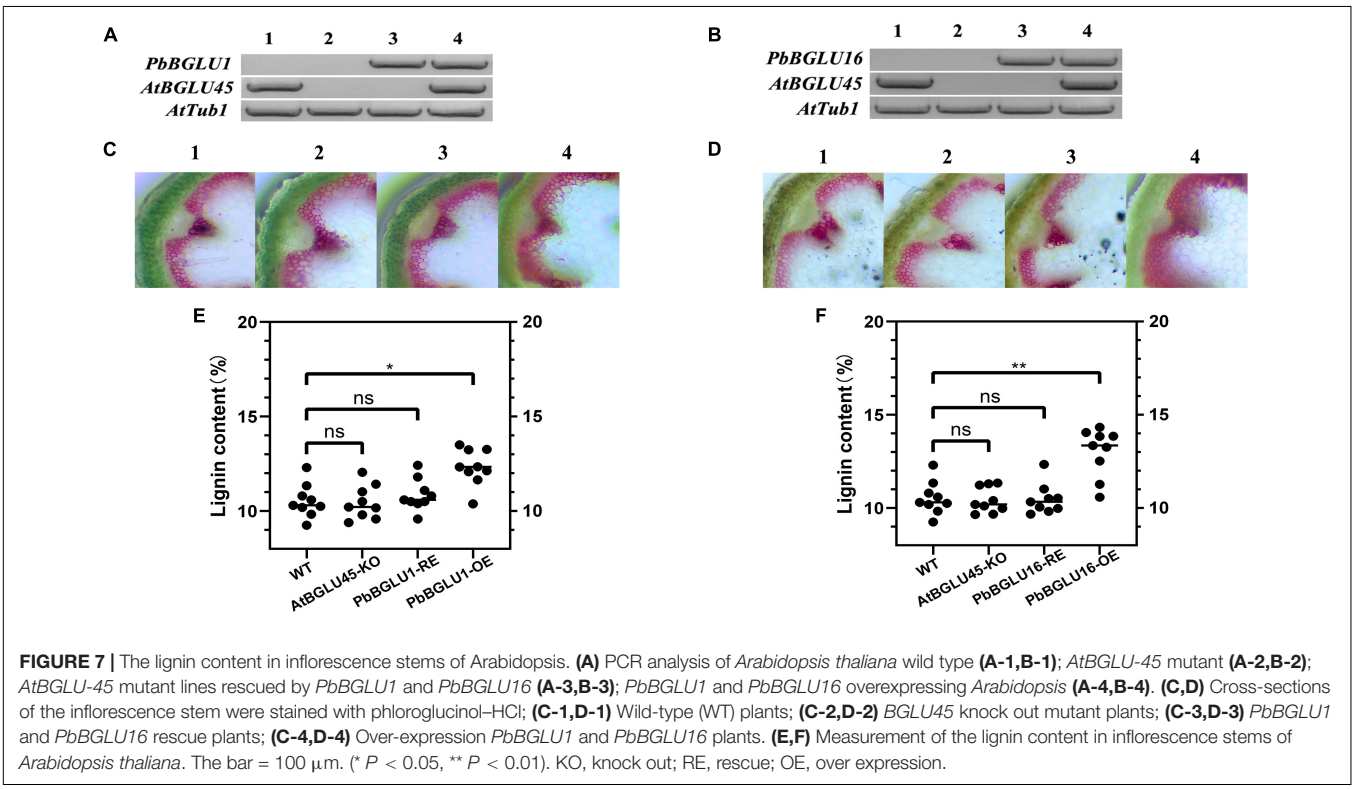


FIGURE 7 | The lignin content in inflorescence stems of Arabidopsis. **(A)** PCR analysis of *Arabidopsis thaliana* wild type **(A-1,B-1)**; *AtBGLU-45* mutant **(A-2,B-2)**; *AtBGLU-45* mutant lines rescued by *PbBGLU1* and *PbBGLU16* **(A-3,B-3)**; *PbBGLU1* and *PbBGLU16* overexpressing *Arabidopsis* **(A-4,B-4)**. **(C,D)** Cross-sections of the inflorescence stem were stained with phloroglucinol–HCl; **(C-1,D-1)** Wild-type (WT) plants; **(C-2,D-2)** *BGLU45* knock out mutant plants; **(C-3,D-3)** *PbBGLU1* and *PbBGLU16* rescue plants; **(C-4,D-4)** Over-expression *PbBGLU1* and *PbBGLU16* plants. **(E,F)** Measurement of the lignin content in inflorescence stems of *Arabidopsis thaliana*. The bar = 100 μm. (* $P < 0.05$, ** $P < 0.01$). KO, knock out; RE, rescue; OE, over expression.

after empty vector, overexpression, and RNAi injection. The results are shown in **Figures 8B,D** which indicate that when *PbBGLU1* and *PbBGLU16* were silenced, the contents of stone cells and lignin in pear fruit did not change compared with empty vector injection. However, in pear fruits with transient overexpression of *PbBGLU1*, the contents of lignin and stone cells in pear fruits were significantly higher ($0.01 < P < 0.05$) than those in pear fruits with empty vector injection. After transient expression of *PbBGLU16*, lignin in pear fruit increased significantly compared with the control group ($0.01 < P < 0.05$), while the content of stone cells in pear fruit injected with *PbBGLU16* showed a very significant difference ($P < 0.01$) compared with the control group.

DISCUSSION

In plants, glycoside hydrolase (GH) family 1 β-glycosidases are believed to play an important role in many diverse processes, including chemical defense against herbivory, hydrolysis of cell wall-derived oligosaccharides during germination, control of active phytohormone levels, and lignification (Xu et al., 2004; Lee et al., 2006; Ahn et al., 2010; Xu et al., 2012; Zhou et al., 2012; Miyahara et al., 2013).

The complete sets of members of the GH1 BGLU genes family have been identified only in a few species. We identified

TABLE 2 | HPLC analysis of soluble monolignol glucosides metabolites in *Arabidopsis*.

Materials	Genotypes	Coniferin (μmol/5 g)	Syringin (μmol/5 g)
Inflorescence stems	WT	0.135 ± 0.03	0.062 ± 0.02
	BGLU45-KO	0.251 ± 0.01	0.104 ± 0.02
	PbBGLU1-RE	0.105 ± 0.01	0.051 ± 0.01
	PbBGLU16-RE	0.126 ± 0.04	0.058 ± 0.03
	PbBGLU1-OE	0.132 ± 0.02	0.059 ± 0.03
	PbBGLU16-OE	0.138 ± 0.03	0.063 ± 0.03

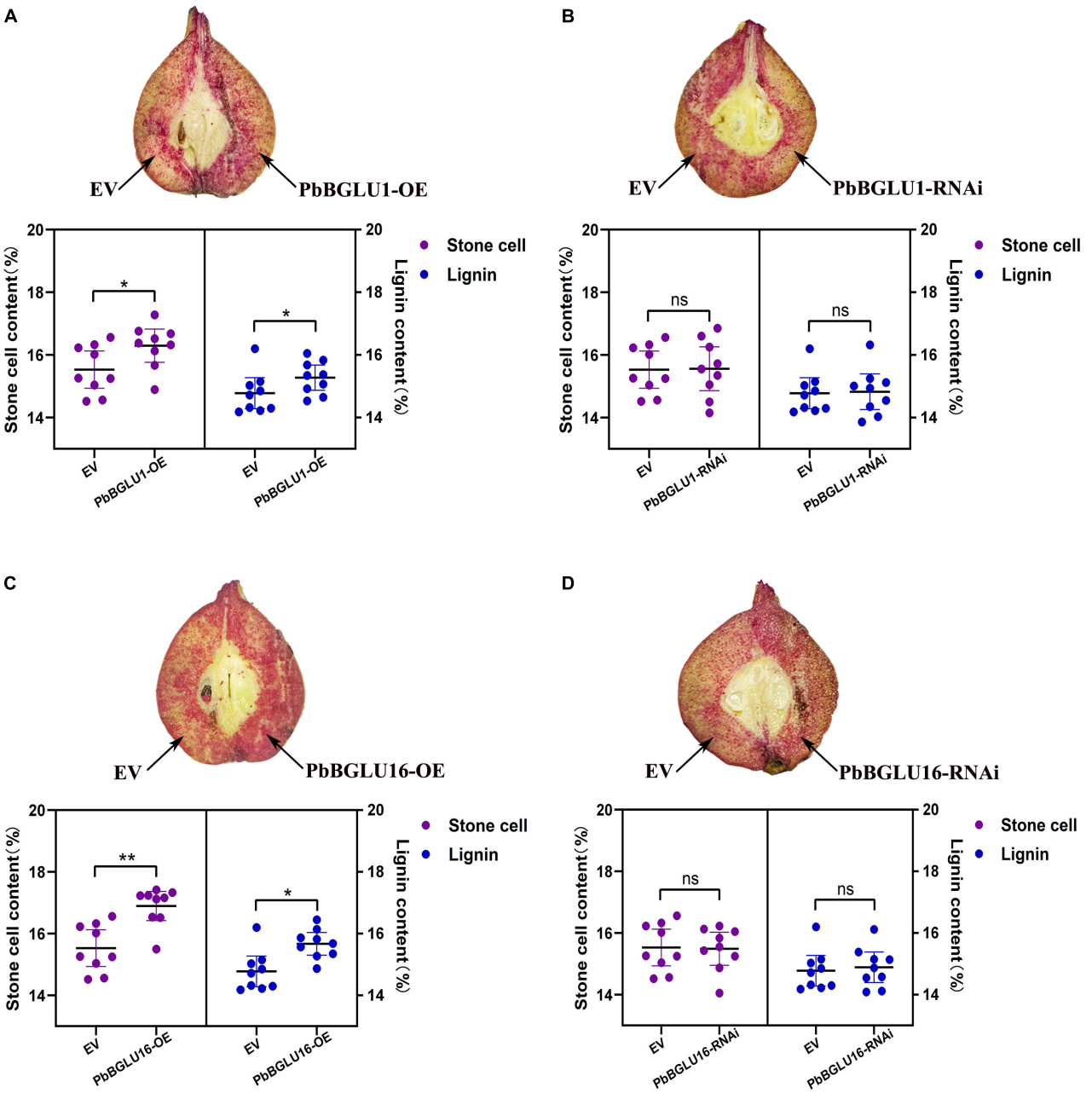


FIGURE 8 | Transient assays using PbBGLU1 and PbBGLU16 silencing and overexpression constructs in "Dangshansuli" fruit at 39 days after flowering (DAF). Images were taken at 10 days after agro-infiltration. Cross-sections of PhBGLU1, PbBGL.U16 overexpression **(A,C)** and silencing **(B,D)** pear fruits were stained with phloroglucinol–HCl. Lignin and stone cell contents in PbBGLU1, PbBGLU16 overexpression **(A,C)** and silencing **(B,D)** flesh tissue around the infiltration sites, more than three fruits were injected with each construct in an experiment that was repeated three times. EV, empty vector; (* $P < 0.05$, ** $P < 0.01$).

a total 50 *BGLUs* in pear. The number was similar to that in *Arabidopsis* (48) (Xu et al., 2004) and poplar (48) (Tsuyama and Takabe, 2015), but more than that in rice (34) (Opassiri et al., 2006) and maize (26) (Gomez-Anduro et al., 2011). Previous studies have shown that the role of monolignol glucosides was necessary for transmembrane transport (Tsuyama and Takabe, 2015; Julien et al., 2016), monolignol glucosides form can increase the solubility and stability of monolignols, which is very beneficial for the transportation of lignin, ultimately to specific locations at the cell wall, after which point lignin begins to deglycosylate and polymerize (Liu, 2012). Many previous studies have shown that BGLU genes are localized to the cell wall and extracellular space (Dharmawardhana et al., 1995; Chapelle et al., 2012; Baiya et al., 2018). In this study, we found that green fluorescence signals from the expressed PbBGLU1-GFP, PbBGLU15-GFP, and PbBGLU16-GFP fusion constructs were specifically distributed

in the extracellular region, which is consistent with previous research. Also, in our *in situ* hybridization analysis, it found that *PbBGLU1* and *PbBGLU16* transcripts were mainly located in lignin deposition and stone cell areas of pear fruit. This result further suggests that *PbBGLU1* and *PbBGLU16* may be functional genes in the process of lignin synthesis and stone cell development in pear.

The heterologous expression of *Arabidopsis* BGLU45 and BGLU46 genes in *Pichia pastoris* revealed that BGLU45 was highly specific for the three monolignol glucosides (coniferin, syringin, and p-coumaryl alcohol 4-O-b-D-glucoside). By contrast, BGLU46 has p-coumaryl alcohol glucoside as the preferred substrate, but it displays a broad specificity toward various substrates (Xu et al., 2004; Escamilla-Treviño et al., 2006). The *in vitro* substrate specificity of BGLU45 and BGLU46 and the expression of their genes in stems (Escamilla-Treviño et al., 2006) support the hypothesis that these enzymes are involved in the lignification of *Arabidopsis*. To further, verify the functions of *PbBGLU1* and *PbBGLU16*, we induced expression and purification recombinant protein, and selected coniferin and syringin as substrates (Dicotyledons mainly contain syringic units and guaiacyl unit lignin: G-S). Enzyme activity analysis showed that the two recombinant proteins could act on coniferin and syringin to deglycosylate them and produce coniferyl alcohol and sinapyl alcohol. Moreover, compared with GST-PbBGLU1 recombinant protein, GST-PbBGLU16 recombinant protein shows higher activity and catalytic efficiency for the two substrates.

In previous studies, the implication of monolignol glucosides in the lignification process has been repeatedly suggested (Boerjan et al., 2003). The hypothesis was that monolignol glucosides would be stored in the vacuole or transported from the cytosol to the cell. One of the most predominant questions about phenylpropanoid biosynthesis concerns the spatial distribution of both precursors and final products within the cells. Many phenylpropanoid molecules necessary for plant development and/or defense are also toxic and an important challenge for plants is how to successfully manage the production/storage of these potentially dangerous molecules (Väisänen et al., 2015). In previous studies, the knock-out of *BGLU-45* and *BGLU-46* in *Arabidopsis* can lead to a significant increase of monolignol glycosides, but the lignin content does not seem to change significantly (Chapelle et al., 2012; Baiya et al., 2018). We also obtained the same results in our experiment. After transferred *PbBGLU1* and *PbBGLU16* into *BGLU*-45 *Arabidopsis* mutant, we found that the rise of these two glycosides in *BGLU*-45 *Arabidopsis* mutant was some degree of recovery, which further proved the function of *PbBGLU1* and *PbBGLU16*. When *PbBGLU1* and *PbBGLU16* were overexpressed into *Arabidopsis*, the lignin content of *A. thaliana* was significantly increased, and the increasing degree of *PbBGLU16-OE* was greater than that of *PbBGLU1-OE*. The same result also occurred in the experiment of instantaneous injection in pear fruit, when we silenced *PbBGLU1* and *PbBGLU16* by RNA interference, there was no significant difference in the contents of lignin and stone cells between pear and the control group, but the contents of stone cells and lignin were significantly increased when *PbBGLU1* and *PbBGLU16* were transiently overexpressed, and PbBGLU16 has a greater effect on lignin and stone cells than PbBGLU1. A past study also showed that the expression of *BGLU45* and *BGLU46* was deregulated under various biological stresses, suggesting that the monolignol storage pool might be used under pathogen attack in the fabrication of "defense lignin" and/or phytoalexins, cell, and vacuole disruption following pathogen attack, insect/herbivore feeding and potentially severe abiotic stress (cell freezing, desiccation) or during programmed cell death (xylem vessel) would lead to contact between themonolignol glucosides and β-glucosidases at the cell wall where PRXs and/or LACs could rapidly polymerize the aglycone forms (Morant et al., 2008; Chapelle et al., 2012). This also proves β-glucosidase can directly or indirectly increase lignin. However, monolignol glucosides had not been considered a dead-end product and its content is still regulated by many parties. Monolignol glucosides had often been detected in gymnosperms and its content in the differentiating xylems of gymnosperms peaks at around the cambium and decreases as lignification progresses (Fukushima et al., 1997; Tsuyama and Takabe, 2015). The regulation mechanism of lignin transport is very complex, which needs to be explored continuously. In conclusion, this study screened and proved the function of β-glucosidase during the process of lignin synthesis and stone cells development in pear fruit, and laid a foundation for improving pear fruit quality at the molecular level.

CONCLUSION

In order to analyze the BGLU β-glucosidases in the glycoside hydrolase family 1 (GH1) of pear, bioinformatics analysis, qRT-PCR, subcellular localization, *in situ* hybridization, enzyme activity analysis, the transformation of *PbBGLU1* and *PbBGLU16* into *Arabidopsis,* and transient injection *PbBGLU1* and *PbBGLU16* into pear fruit experiment were performed in this study. A total of 50 non-redundant BGLU family members were identified in Chinese white pears while three candidate genes *PbBGLU1, PbBGLU15,* and *PbBGLU16* were identified that might be involved in lignin synthesis. Subcellular localization showed that all three genes were located on the extracellular region. *In situ* hybridization analysis revealed that *PbBGLU1* and *PbBGLU16* were involved in lignin synthesis and stone cell development in pear fruit. Enzyme activity analysis showed that GST-PbBGLU1 and GST-PbBGLU16 had activity on coniferin and syringin. The transformation of *PbBGLU1* and *PbBGLU16* into *Arabidopsis* demonstrated that as compare with wild-type *Arabidopsis,* the overexpression plants lignin increased in varying degrees and the effect of *PbBGLU16* on the increase of lignin in *Arabidopsis* was greater than that of *PbBGLU1*. In pear fruits with transient overexpression of *PbBGLU1* and *PbBGLU16*, the contents of *PbBGLU1-OE* and *PbBGLU16-OE* lignin and stone cells in pear fruits were significantly higher than those in pear fruits with empty vector injection, the effect of *PbBGLU16* on the increase of lignin and stone cells was greater than that of *PbBGLU1*. In summary, the identification

of 50 *BGLUs* and analysis of these two BGLUs gene functions in pear will help to improve the quality of pear fruit in the future. These results provide information that may facilitate further functional analyses of PbBGLU genes to elucidate their biological roles in pear.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YoC conceived and designed the experiments. HW, YiZ, and XF performed the experiments. FP, MM, YaZ, WH, YuZ, JL, and YuC analyzed the data. HW wrote the manuscript. All authors reviewed and approved the final submission.

FUNDING

This study was supported by National Natural Science Foundation of China (No. 31640068), Anhui Agricultural University 2020 Graduate Innovation Foundation (2020ysj-50), and 2021 Postgraduate Research Project in Colleges and Universities in Anhui Province (YJS20210202).

ACKNOWLEDGMENTS

We would like to thank everyone who contributed to this article. This manuscript part of the data has been released as a pre-print at Research Square (Wang et al., 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.852001/full#supplementary-material

REFERENCES

Ahn, Y. O., Shimizu, B., Sakata, K., Gantulga, D., Zhou, C., Esen, A., et al. (2010). Scopolin-hydrolyzing beta-glucosidases in roots of *Arabidopsis*. *Plant Cell Physiol.* 51, 132–143. doi: 10.1093/pcp/pcp174

Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Martinoia, E., et al. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Curr. Biol.* 22, 1207–1212. doi: 10.1016/j.cub.2012.04.064

Anderson, N. A., Tobimatsu, Y., Ciesielski, P. N., Ximenes, E., Ralph, J., Donohoe, B. S., et al. (2015). Manipulation of guaiacyl and syringyl monomer biosynthesis in an *Arabidopsis* cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell* 27, 2195–2209. doi: 10.1105/tpc.15.00373

Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME suite. *Nucleic Acids Res.* 43, W39–W49. doi: 10.1093/nar/gkv416

Baiya, S., Hua, Y., Ekkhara, W., and Ketudat Cairns, J. R. (2014). Expression and enzymatic properties of rice (*sativa* L.) monolignol β-glucosidases. *Plant Sci.* 227, 101–109. doi: 10.1016/j.plantsci.2014.07.009

Baiya, S., Mahong, B., Lee, S. K., Jeon, J. S., and Ketudat Cairns, J. R. (2018). Demonstration of monolignol beta-glucosidase activity of rice Os4BGlu14, Os4BGlu16 and Os4BGlu18 in *Arabidopsis thaliana* bglu45 mutant. *Plant Physiol. Biochem.* 127, 223–230. doi: 10.1016/j.plaphy.2018.03.026

Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot.* 15, 1053–1074. doi: 10.1093/aob/mcv046

Biely, P., Ahlgren, J. A., Leathers, T. D., Greene, R. V., and Cotta, M. A. (2003). Aryl-glycosidase activities in germinating maize. *Cereal Chem.* 80, 144–147. doi: 10.1094/CCHEM.2003.80.2.144

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

Cai, Y., Li, G., Nie, J., Lin, Y., Nie, F., Xu, Y., et al. (2010). Study of the structure and biosynthetic pathway of lignin in stone cells of pear. *Sci. Hortic.* 125, 374–379. doi: 10.1016/j.scienta.2010.04.029

Cao, Y., Han, Y., Li, D., Lin, Y., and Cai, Y. (2016a). MYB transcription factors in chinese pear (*Pyrus bretschneideri* Rehd.): genome-wide identification, classification, and expression profiling during fruit development. *Front. Plant Sci.* 7:577. doi: 10.3389/fpls.2016.00577

Cao, Y., Li, D. H., Meng, D. D., Jin, Q., Cai, Y. P., and Lin, Y. (2016b). Structural, evolutionary, and functional analysis of the Class III peroxidase gene family in chinese pear (*Pyrus bretschneideri*). *Front. Plant Sci.* 7:1874. doi: 10.3389/fpls.2016.01874

Cao, Y., Han, Y., Meng, D., Li, D., Jiao, C., Cai, Y., et al. (2017). B-BOX genes: genome-wide identification, evolution and their contribution to pollen growth in pear (*Pyrus bretschneideri* Rehd.). *BMC Plant Biol.* 17:156. doi: 10.1186/s12870-017-1105-4

Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Demont-Caulet, N., et al. (2012). Impact of the absence of stem-specific beta-glucosidases on lignin and monolignols. *Plant Physiol.* 160, 1204–1217. doi: 10.1104/pp.112.203364

Cheng, X., Li, G., Manzoor, M. A., Wang, H., Abdullah, M., Lin, Y., et al. (2019a). In silico genome-wide analysis of respiratory burst oxidase homolog (RBOH) family genes in five fruit-producing trees, and potential functional analysis on lignification of stone cells in Chinese White pear. *Cells* 8:520. doi: 10.3390/cells8060520

Cheng, X., Muhammad, A., Li, G., Zhang, J., Cheng, J., Lin, Y., et al. (2019b). Family-1 UDP glycosyltransferases in pear (*Pyrus bretschneideri*): molecular identification, phylogenomic characterization and expression profiling during stone cell formation. *Mol. Biol. Rep.* 46, 2153–2175. doi: 10.1007/s11033-019-04669-y

Dharmawardhana, D. P., Ellis, B. E., and Carlson, J. E. (1995). A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. *Plant Physiol.* 107, 331–339. doi: 10.2307/4276327

Dharmawardhana, D. P., Ellis, B. E., and Carlson, J. E. (1999). cDNA cloning and heterologous expression of coniferin β-glucosidase. *Plant Mol. Biol.* 40, 365–372. doi: 10.1023/A:1006226931512

Dong, X., Jiang, Y., and Hur, Y. (2019). Genome-wide analysis of glycoside hydrolase family 1 beta-glucosidase Genes in *Brassica rapa* and their potential role in pollen development. *Int. J. Mol. Sci.* 20:1663. doi: 10.3390/ijms20071663

Eastwood, D. C., Challen, M. P., Zhang, C., Jenkins, H., Henderson, J., and Burton, K. S. (2008). Hairpin-mediated down-regulation of the urea cycle enzyme argininosuccinatelyase in *Agaricus bisporus*. *Mycol. Res.* 112(Pt 6), 708–716. doi: 10.1016/j.mycres.2008.01.009

Escamilla-Treviño, L. L., Chen, W., Card, M. L., Shih, M. C., Cheng, C. L., Poulton, J. E., et al. (2006). *Arabidopsis thaliana* beta-Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. *Phytochemistry* 67, 1651–1660. doi: 10.1016/j.phytochem.2006.05.022

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Bateman, A., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211–D222. doi: 10.1093/nar/gkp985

Fukushima, K., Taguchi, S., Matsui, N., and Yasuda, S. (1997). Distribution and seasonal changes of monolignol glucosides in *Pinus thunbergii*. *Mokuzai Gakkaishi* 43, 254–259.

Gachon, C. M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. *Trends Plant Sci*. 10, 542–549. doi: 10.1016/j.tplants.2005.09.007

Gomez-Anduro, G., Ceniceros-Ojeda, E. A., Casados-Vazquez, L. E., Bencivenni, C., Sierra-Beltran, A., Tiessen, A., et al. (2011). Genome-wide analysis of the beta-glucosidase gene family in maize (*Zea mays* L. var B73). *Plant Mol. Biol*. 77, 159–183. doi: 10.1007/s11103-011-9800-2

Heneghan, M. N., Costa, A. M. S. B., Challen, M. P., Peter, R. M., Bailey, A., and Foste, G. D. (2007). A comparison of methods for successful triggering of gene silencing in *Coprinopsis cinerea*. *Mol. Biotechnol*. 35, 283–296. doi: 10.1007/BF02686014

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., Gao, G., et al. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Ishihara, H., Tohge, T., Viehover, P., Fernie, A. R., Weisshaar, B., Stracke, R., et al. (2016). Natural variation in flavonol accumulation in *Arabidopsis* is determined by the flavonol glucosyltransferase BGLU6. *J. Exp. Bot*. 67, 1505–1517. doi: 10.1093/jxb/erv546

Jin, Q., Yan, C., Qiu, J., Zhang, N., Lin, Y., and Cai, Y. (2013). Structural characterization and deposition of stone cell lignin in Dangshan Su pear. *Sci. Hortic*. 155, 123–130. doi: 10.1016/j.scienta.2013.03.020

Julien, L. R., Brigitte, H., Anne, C., Simon, H., and Godfrey, N. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci*. 7:735. doi: 10.3389/fpls.2016.00735

Kristoffersen, P., Brzobohaty, B., Höhfeld, I., Bako, L., Melkonian, M., and Palme, K. (2000). Developmental regulation of the maize Zm-p60.1 gene encoding a beta-glucosidase located to plastids. *Planta* 210, 407–415. doi: 10.1007/PL00008149

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. 33, 1870–1874. doi: 10.1093/molbev/msw054

Lanot, A., Hodge, D., Jackson, R. G., George, G. L., Elias, L., Lim, E. K., et al. (2006). The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in *Arabidopsis thaliana*. *Plant J*. 48, 286–295. doi: 10.1111/j.1365-313X.2006.02872.x

Lee, K. H., Piao, H. L., Kim, H. Y., Choi, S. M., Jiang, F., Hwang, I., et al. (2006). Activation of glucosidase *via* stress-induced polymerization rapidly increases active pools of abscisic acid. *Cell* 126, 1109–1120. doi: 10.1016/j.cell.2006.07.034

Leinhos, V., Udagama-Randeniya, P. V., and Savidge, R. A. (1994). Purification of an acidic coniferin-hydrolysing beta-glucosidase from developing xylem of *Pinus banksiana*. *Phytochemistry* 37, 11–15. doi: 10.1016/0031-9422(94)85053-4

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res*. 40, D302–D305. doi: 10.1093/nar/gkr931

Liu, C. J. (2012). Deciphering the enigma of lignification: precursor transport, oxidation, and the topochemistry of lignin assembly. *Mol. Plant* 5, 304–317. doi: 10.1093/mp/ssr121

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

Miao, Y. C., and Liu, C. J. (2010). ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. *Proc. Natl. Acad. Sci. U.S. A*. 107, 22728–22733. doi: 10.1073/pnas.1007747108

Miyahara, T., Sakiyama, R., Ozeki, Y., and Sasaki, N. (2013). Acyl-glucose-dependent glucosyltransferase catalyzes the final step of anthocyanin formation in *Arabidopsis*. *J. Plant Physiol*. 170, 619–624. doi: 10.1016/j.jplph.2012.12.001

Morant, A. V., Jorgensen, K., Jorgensen, C., Paquette, S. M., SanchezPerez, R., Moller, B. L., et al. (2008). beta-Glucosidases as detonators of plant chemical defense. *Phytochemistry* 69, 1795–1813. doi: 10.1016/j.phytochem.2008.03.006

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J. R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 beta-glucosidase. *BMC Plant Biol*. 6:33.

Perkins, M., Smith, R. A., and Samuels, L. (2019). The transport of monomers during lignification in plants: anything goes but how? *Curr. Opin. Biotechnol*. 56, 69–74. doi: 10.1016/j.copbio.2018.09.011

Pradhan, M. P., and Loqué, D. (2014). Histochemical staining of *Arabidopsis thaliana* secondary cell wall elements. *J. Vis. Exp*. 87:51381. doi: 10.3791/51381

Raiskila, S., Pulkkinen, M., Laakso, T., Fagerstedt, K., Löija, M., Mahlberg, R., et al. (2007). FTIR spectroscopic prediction of Klason and acid soluble lignin variation in Norway spruce cutting clones. *Silva Fenn*. 41, 351–371. doi: 10.14214/sf.301

Roepke, J., Gordon, H. O. W., Neil, K. J. A., Gidda, S., Mullen, R. T., Freixas Coutin, J. A., et al. (2017). An apoplastic β-glucosidase is essential for the degradation of flavonol 3-O-β-glucoside-7-O-α-rhamnosides in *Arabidopsis*. *Plant Cell Physiol*. 58, 1030–1047. doi: 10.1093/pcp/pcx050

Steeves, V., Förster, H., Pommer, U., and Savidge, R. (2001). Coniferyl alcohol metabolism in conifers – I. Glucosidic turnover of cinnamyl aldehydes by UDPG: coniferyl alcohol glucosyltransferase from pine cambium. *Phytochemistry* 57, 1085–1093. doi: 10.1016/S0031-9422(01)00107-8

Su, X., Zhao, Y., Wang, H., Li, G., Cheng, X., Cai, Y., et al. (2019). Transcriptomic analysis of early fruit development in Chinese white pear (*Pyrus bretschneideri* Rehd.) and functional identification of PbCCR1 in lignin biosynthesis. *BMC Plant Biol*. 19:417. doi: 10.1186/s12870-019-2046-x

Tsuji, Y., and Fukushima, K. (2004). Behavior of monolignol glucosides in angiosperms. *J. Agric. Food Chem*. 52, 7651–7659. doi: 10.1021/jf0491198

Tsuyama, T., and Takabe, K. (2015). Coniferin β-glucosidase is ionically bound to cell wall in differentiating xylem of poplar. *J. Wood Sci*. 61, 438–444. doi: 10.1007/s10086-015-1486-7

Väisänen, E. E., Smeds, A. I., Fagerstedt, K. V., Teeri, T. H., Willför, S. M., and Kärkönen, A. (2015). Coniferyl alcohol hinders the growth of tobacco BY-2 cells and *Nicotiana benthamiana* seedlings. *Planta* 242, 747–760. doi: 10.1007/s00425-015-2348-7

Wang, H., Zhang, Y., Zhao, Y., Han, W., Lu, J., Cai, Y., et al. (2020). Genome-wide comparative analysis of the β-glucosidase family in five rosaceae species and their potential role on lignification of stone cells in Chinese white Pear. *Res. Squ.* [Preprint]. doi: 10.21203/rs.3.rs-46971/v1

Wu, J., Wang, Z., Shi, Z., Zhang, S., Ming, R., Zhang, S., et al. (2012). The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res*. 23, 396–408. doi: 10.1101/gr.144311.112

Wu, M. F., and Wagner, D. (2012). "RNA in situ hybridization in *Arabidopsis*," in *RNA Abundance Analysis. Methods in Molecular Biology (Methods and Protocols)*, Vol. 883, eds H. Jin and W. Gassmann (Totowa, NJ: Humana Press).

Wu, T., Zhang, C., Wu, J., Wan, H., and Zhang, S. (2012). Evaluation of candidate reference genes for real time quantitative PCR normalization in pear fruit. *Afr. J. Agric. Res*. 7, 3701–3704. doi: 10.5897/AJAR11.1842

Xie, M., Huang, Y., Zhang, Y., Wang, X., Yang, H., Yu, O., et al. (2013). Transcriptome profiling of fruit development and maturation in chinese white pear (*Pyrus bretschneideri* rehd). *BMC Genomics* 14:823. doi: 10.1186/1471-2164-14-823

Xu, Z. Y., Lee, K. H., Dong, T., Jeong, J. C., Jin, J. B., Hwang, I., et al. (2012). A vacuolar beta-glucosidase homolog that possesses glucose-conjugated abscisic acid hydrolyzing activity plays an important role in osmotic stress responses in *Arabidopsis*. *Plant Cell* 24, 2184–2199. doi: 10.1105/tpc.112.095935

Xu, Z., Luis, L., Zeng, L. H., Mallikarjun, L., David, R., Cheng, C. L., et al. (2004). Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. *Plant Mol. Biol*. 55, 343–367. doi: 10.1007/s11103-004-0790-1

Yan, C., Yin, M., Zhang, N., Jin, Q., Fang, Z., Cai, Y., et al. (2014). Stone cell distribution and lignin structure in various pear varieties. *Sci. Hortic*. 174, 142–150. doi: 10.1016/j.scienta.2014.05.018

Zamioudis, C., Hanson, J., and Pieterse, C. M. J. (2014). β-glucosidase bglu42 is a myb72-dependent key regulator of rhizobacteria-induced systemic resistance and modulates iron deficiency responses in *Arabidopsis* roots. *New Phytol*. 204, 368–379. doi: 10.1111/nph.12980

Zhang, J., Cheng, X., Jin, Q., Su, X., Li, M., Cai, Y., et al. (2017). Comparison of the transcriptomic analysis between two Chinese white pear (*Pyrus bretschneideri* Rehd.) genotypes of different stone cells contents. *PLoS One* 12:e0187114. doi: 10.1371/journal.pone.0187114

Zhang, M. Y., Xue, C., Hu, H., Li, J., Xue, Y., Wang, R., et al. (2021). Genome-wide association studies provide insights into the genetic determination of fruit traits of pear. *Nat. Commun*. 12:1144. doi: 10.1038/s41467-021-21378-y

Zhou, C., Tokuhisa, J. G., Bevan, D. R., and Esen, A. (2012). Properties of beta-thioglucoside hydrolases (TGG1 and TGG2) from leaves of *Arabidopsis thaliana*. *Plant Sci.* 191, 82–92. doi: 10.1016/j.plantsci.2012.02.004

Zhou, H., Qi, K., Xing, L., Hao, Y., Peng, W., Zhang, S., et al. (2016). Genome-wide identification and comparative analysis of the cation proton antiporters family in pear and four other Rosaceae species. *Mol. Gen. Genom*. 291:1727.

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fpls.2022.853444

Depicting Precise Temperature and Duration of Vernalization and Inhibiting Early Bolting and Flowering of *Angelica sinensis* by Freezing Storage

Xiaoxia Liu[1†], Mimi Luo[1†], Mengfei Li[1,2] and Jianhe Wei[2*]*

[1] State Key Laboratory of Aridland Crop Science, College of Life Science and Technology, Gansu Agricultural University, Lanzhou, China, [2] Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Einat Shemesh Mayer,
Agricultural Research Organization (ARO), Israel
Wenxuan Dong,
Shenyang Agricultural University, China

***Correspondence:**
Mengfei Li
lmf@gsau.edu.cn
Jianhe Wei
jhwei@implad.ac.cn

[†] These authors have contributed equally to this work

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

Received: 12 January 2022
Accepted: 19 April 2022
Published: 19 May 2022

Citation:
Liu X, Luo M, Li M and Wei J (2022) Depicting Precise Temperature and Duration of Vernalization and Inhibiting Early Bolting and Flowering of Angelica sinensis by Freezing Storage. Front. Plant Sci. 13:853444. doi: 10.3389/fpls.2022.853444

Angelica sinensis is a perennial rhizomatous herb that is widely used for the treatment of cardio-cerebrovascular diseases, which largely rely on metabolites, such as alkylphthalides, polysaccharides, and ferulic acid. This plant must experience low-temperature vernalization and long-day conditions for the occurrence of early bolting and flowering (EBF) that reduces yield and quality of fleshy root. In current commercial planting, the EBF of more than 40% is mainly attributed to the completion of vernalization of seedlings during overwinter storage. While effects of storage temperatures [vernalization temperature (0–10°C) and freezing temperature (−2 to −12°C)] and seedling sizes on the EBF have been observed in previous studies, the precise vernalization temperature and duration for different size seedlings, the effective freezing storage to avoid vernalization of seedlings, and physiological characteristics have not been systematically investigated. Here, the EBF rate, the anatomical structure of shoot apical meristem (SAM), and physiological characteristics of different size seedlings at different storage temperatures (0, 3, 5, −3, and −5°C) and durations (14–125 d) are reported. The vernalization duration of seedlings was predicated from 57 to 85 d with temperatures ranging from 0 to 5°C based on the linearization regression analysis *via* Matrix Laboratory software. The EBF can be effectively inhibited by freezing storage. The anatomical structure of SAM, levels of primary metabolites (soluble sugar, starch, amino acid, and protein), and endogenous hormones (GA_3, IAA, and ABA), exhibited a dynamical change in the seedlings at different storage temperatures. These findings will provide useful information for predicting the vernalization of seedlings and inhibiting the EBF in large-scale commercial cultivation.

Keywords: *Angelica sinensis* (Oliv.) Diels, vernalization, early bolting and flowering, freezing storage, physiological characteristics

Abbreviations: EBF, early bolting and flowering; FIR, freezing injury rate; FT, freezing temperature; GR, germination rate; RR, rotting rate; SAM, shoot apical meristem; SC, seedling classification; VD, vernalization duration; VT, vernalization temperature.

INTRODUCTION

Angelica sinensis (Oliv.) Diels (family Umbelliferae), common names as Danggui, Dong quai, Tang kuei, and Chinese angelica, is a perennial herbaceous species (McGuffin et al., 2000). It prefers growing in cool-moist conditions at an altitude of 2,200 to 3,000 m and is widely cultivated in western parts of China, including Gansu, Qinghai, Sichuan, and Yunnan (Zhang and Cheng, 1989; Zhang et al., 2012; Xu et al., 2020). The roots have been used as a traditional Chinese medicine for nourishing and activating the blood, regulating female menstrual disorders and relieving pains, relaxing bowels, etc., over 2,000 years (Upton, 2003; Committee for the Pharmacopoeia of PR China, 2015; Wei et al., 2016). In recent years, the roots are also been applied in the treatment of cardio-cerebrovascular diseases as well as agents of anti-inflammatory and antioxidant (Upton, 2003; Wang and Ou-Yang, 2005; Chao et al., 2010; Wang et al., 2015), which largely rely on the bioactive components, including alkylphthalides, polysaccharides, ferulic acid, and essential oils (Ma et al., 2015; Wei et al., 2016; Li et al., 2022).

Currently, the cultivated area of *A. sinensis* is more than 43,500 ha due to increasing demand for clinical application (Zhang et al., 2012; Huang and Jin, 2018). In the commercial large-scale cultivation, seeds are sown in summer and germinated seedlings are collected in autumn to be overwintered indoors; in the spring, stored seedlings are planted out for vegetative growth and either harvested in autumn of this second year for fleshy roots or kept in the field until mid-summer of the third year for seed collection (**Supplementary Figure 1**; Huang and Jin, 2018; Li et al., 2021). However, the 2-year-old plants occur up to 40% EBF, which makes the roots lignified along with the reduction in yield and quality (**Supplementary Figure 2**); furthermore, the lignified roots are useless in medicinal agents due to little accumulation of the bioactive components (Zhang et al., 2018; Li et al., 2020a,b, 2022).

To inhibit the occurrence of the EBF, several efforts being made include selecting the cultivars with a lower rate of EBF (Huang and Jin, 2018), controlling the seedling sizes (Lin et al., 2007), investigating the type of vernalization and photoperiod (Wang, 1977), storing the seedlings below freezing temperatures (Wang, 1977, 1979; Jia et al., 2017, 2018), and avoiding the plants grown in the long-day conditions (Yao, 2005). Specifically, Huang and Jin (2018) reported that the EBF rate of the green stem cultivar (Mingui 2) was lower than the purple stem cultivar (Mingui 1), while the Mingui 1 occupied over 95% cultivation area due to better growth characteristics. Lin et al. (2007) demonstrated that there was a positive relationship of EBF rates with seedling sizes, with the EBF rates of 73, 47, 25, and 6% under the root shoulder diameter $\geq$0.66, 0.56–0.65, 0.46–0.55, and $\leq$0.45 cm, respectively. Wang (1977) found that *A. sinensis* is a "low-temperature and long-day" plant and the transition from vegetative growth to flowering must satisfy with vernalization (0–5°C) and long-day conditions (>12 h daylight). Based on the type of vernalization and photoperiod, the vernalization could be avoided after the seedlings are stored at −2 to −12°C (Wang, 1979; Jia et al., 2018); the long-day conditions could be also avoided under 40, 60, and 80% sunshade with the EBF rates 16.3, 12.3, and 5.3%, respectively (Yao, 2005). In the practical large-scale cultivation to inhibit or reduce the EBF, avoiding the vernalization of seedlings stored in a smaller freezing room is far more feasible than shading the long-day conditions of plants grown in a larger sunshade field (Wang, 1977).

Although effective temperature range and duration for vernalization of *A. sinensis* seedlings have been reported in previous literature, their results for vernalization temperature (VT) and duration (VD) are inconsistent (Wang, 1977; Li, 1979; Zhang and Huang, 1998). Specifically, Wang (1977) found that the effective temperature for vernalization ranged from 0 to 5°C; Li (1979) reported that more than 80°C accumulated temperature was required for flowering with the seedlings exposed to 5–10°C for a certain duration; and Zhang and Huang (1998) reported that the vernalization could complete at 4°C for 170 d. Furthermore, changes in physiological characteristics, such as the levels of soluble sugar, protein, malonic dialdehyde, and nitrate reductase, were observed when the seedlings were stored at different vernalization and freezing temperatures (Zhang and Huang, 1998; Chen et al., 2014). The abovementioned literature indicate that the studies on the precise VT and VD for different seedlings (in other words the most effective conditions to induce EBF), the effective freezing storage to avoid vernalization of seedlings, and physiological characteristics during storage are still limited, which lead to the problem of the EBF having not been resolved so far. In this study, the effects of storage temperatures [vernalization temperature (VT, 0–5°C) and freezing temperature (FT, −3 to −5°C)] and durations (14–125 d) on the rotting or chilling damages, EBF, anatomical structure of SAM, levels of primary metabolites, and endogenous hormones in different size seedlings were systematically investigated by the methods of paraffin section, spectrometry, and HPLC.

MATERIALS AND METHODS

Plant Materials

The mature seeds of 3-year-old *A. sinensis* (cultivar Mingui 1) were sown (3,020 m a.s.l.; 36°59′ 41″N, 102°4′ 29″E) on June 26, 2020. After 108 d germination and growth *in situ*, the seedlings were dug out on October 11, 2020 (**Supplementary Figure 3**). After leaving in outdoor for about 10 d to evaporate water content to 70–75%, the seedlings were stored at different temperatures (0, 3, 5, −3, and −5°C). After certain days (specifically, 0°C for 14, 21, 25, 30, 45, 50, 60, 75, 81, 90, 104, and 125 d; 3 and 5°C for 14, 21, 30, 45, 60, 75, and 90 d; and −3 and −5°C for 25, 50, 81, 104, and 125 d) (**Table 1**), the seedlings were taken out and divided into three grades [large (0.5–0.6 cm), medium (0.4–0.5 cm), and small sizes (0.3–0.4 cm)] based on the diameter of root shoulder measured by a digital caliper, to observe the anatomical structure of SAM, to measure physiological characteristics (e.g., root activity, contents of soluble sugar, starch, protein, and amino acid as well as GA_3, IAA, and ABA), and to investigate the EBF rate with the seedlings planted in a pot (13 cm × 9 cm) with soil (coconut coir: peat: fermented cow

dung: pearlite = 3:3:2:2) and grown at the greenhouse (20°C, 12-h daylight long-day condition).

Survey of Germination and Growth Characteristics

The rotting rate (RR) and freezing injury rate (FIR) of seedlings were immediately surveyed with vernalization (0, 3, and 5°C) at 21, 30, and 45 d and freezing temperatures (−3 and −5°C) at 70, 104, and 125 d; the germination rate (GR) and EBF rate were surveyed after the seedlings grown at 30 and 90 d, respectively. The specific calculations are as follows:

RR (%) = (number of rotten roots of seedlings / n) × 100%, (note: n = 800–1,000)
FIR (%) = (number of frozen injury roots of seedlings/n) × 100%, (note: n = 800–1,000)
GR (%) = (number of germination roots of seedlings n) × 100%, (note: n = 54)
EBF rate (%) = (number of EBF plants/n) × 100%, (note: n = germinated and survived plants, *ca.* 30–50)

Observation of Shoot Apical Meristem

The anatomical structure of shoot apical meristem (SAM) was observed using a paraffin section method with slight modifications (Li et al., 2016, 2020c). Briefly, the roots were first rinsed, and the root shoulders containing the SAM with leaf primordia (0.5 cm) were cut off and immersed into FAA fixative solution (70% ethanol: formaldehyde: glacial acetic acid = 90:5:5, v/v) at 4°C for 12 h. Second, the FAA-fixed samples were washed three times with 70% ethanol at 22°C for 10 min and then sequentially dehydrated in ethanol 30% (4 h), 50% (4 h), 70% (3 h), 85% (3 h), 95% (2.5 h), 100% (2 h), and 100% (1.5 h). Third, the dehydrated samples were sequentially transparentized in the mixture (2:1, 1:1, 1:2, and 0:1 v/v) of ethanol and dimethylbenzene for 2 h, respectively. Fourth, the transparentized samples were sequentially immersed in the mixture (1:1 and 2:1 v/v) of dimethylbenzene and paraffin at 56°C for 12 h, immersed in paraffin three times at 58°C for 12 h, and then embedded in 2 cm paraffin cubes. Finally, the anatomical structure of SAM was observed by a reset one inverted microscope (Revolve RVL-100-G, ECHO, CA, United States) after the embedded samples were sliced (7 μm) with a rotary microtome (KD-2258, Cody, Jinhua, China) and stained with safranin O fast green FCF (S8020, Solarbio, Beijing, China; F8130, Solarbio, Beijing, China).

TABLE 1 | Duration of seedlings stored at different temperatures.

Temperature (°C)	Duration (d)											
0	14	21	25	30	45	50	60	75	81	90	104	125
3	14	21	/	30	45	/	60	75	/	90	/	/
5	14	21	/	30	45	/	60	75	/	90	/	/
-3	/	/	25	/	/	50	/	/	81	/	104	125
-5	/	/	25	/	/	50	/	/	81	/	104	125

The "/" indicates seedlings were not collected.

Measurement of Physiological Characteristics

Measurement of Root Activity

Root activity was measured using a triphenyl tetrazolium chloride (TTC) method (Li and Zhang, 2016). Briefly, the roots were rinsed and cut into small pieces (5 mm). The pieces (0.5 g) were put into 10 mL tubes, and then TTC (3 mL, 0.4% w/v) and phosphate buffer (3 mL, pH 7) were added in sequence. After oscillation and reaction in dark at 37°C for 40 min, sulfuric acid (2 mL, 1 mol/L) was added to stop the reaction. The samples were dried with absorbent paper, grind into homogenate in ethyl acetate (5 mL), and centrifuged at 5,000 r/min at 4°C for 5 min, and then, the supernatant was increased to 9 mL with ethyl acetate. Absorbance was measured at 485 nm, root activity was evaluated based on μg of TTC, and the standard curve of TTC was attached in **Supplementary Table 1**.

Measurement of Soluble Sugar, Starch, Protein, and Amino Acid Contents

Preparation for Extracts

The extracts were prepared according to a published protocol with slight modifications (Yang et al., 2016). Briefly, the freshly collected roots were first air-dried in a ventilated room; the roots were finely ground, and the powder (2.0 g) was soaked in ethanol (15 mL, 10% v/v) and agitated in a shaker with 120 r/min at 24°C for 8 h; the homogenate was centrifuged at 6,000 r/min at 4°C for 10 min and re-extracted twice more. The supernatant was increased to 50 mL with ethanol (10% v/v) and then kept at 4°C for measurement.

Measurement of Soluble Sugar Content

Soluble sugar content was measured using a phenol–sulfuric acid method (Dubois et al., 1956). Extracts (10 μL) were added to phenol reagent (1 mL, 9% v/v), and sulfuric acid (3 mL) was added after oscillation which then reacted at 22°C for 30 min. Absorbance was measured at 485 nm, the soluble sugar content was evaluated based on mg of sucrose, and the standard curve of sucrose was attached in **Supplementary Table 1**.

Measurement of Starch Content

Starch content was measured using an anthrone colorimetry method (Cao et al., 2018). Briefly, of the soluble sugar residues, ddH_2O (15 mL) was added and boiled for 15 min; perchloric acid (2 mL, 9.2 mol/L) was added to the mixture and then boiled for 15 min; the homogenate was centrifuged at 5,000 r/min at 4°C for 10 min. The supernatant was increased to 25 mL with ddH_2O. Extracts (20 μL) were added into ddH_2O (2 mL) and anthrone–sulfuric acid (6 mL) and then boiled for 7 min. Absorbance was measured at 640 nm, starch content was evaluated based on mg of soluble starch, and the standard curve of soluble starch was attached in **Supplementary Table 1**.

Measurement of Protein Content

Protein content was measured using a Coomassie brilliant blue colorimetric method (Bradford, 1976). Briefly, extracts (70 μL) were added into Coomassie brilliant blue G-250 protein reagent

(5 mL) and then reacted at 22°C for 2 min. Absorbance was measured at 595 nm, protein content was evaluated based on mg of bovine serum albumin, and the standard curve of bovine serum albumin was attached in **Supplementary Table 1**.

Measurement of Amino Acid Content

Amino acid content was measured using a ninhydrin coloration method (Wang, 2006). Briefly, extracts (450 μL) were sequentially added into ammonia-free distilled water (1 mL), ninhydrin hydrate (3 mL), and ascorbic acid (0.5 mL); the mixture was put into an 80°C water bath to react for 20 min; after cooling to temperature, the mixture was increased to 20 mL with ammonia-free distilled water. Absorbance was measured at 570 nm, amino acid content was evaluated based on mg of leucine, and the standard curve of leucine was attached in **Supplementary Table 1**.

Quantification of GA_3, IAA, and ABA Contents

The extracts were prepared according to a published protocol with slight modifications (Pan et al., 2010). Briefly, the freshly collected root shoulders containing the SAM with leaf primordia (0.5 g) were first extracted with methanol (7 mL, 80% v/v) at 4°C for 8 h; the homogenate was centrifuged at 8,000 r/min at 4°C for 10 min and re-extracted twice more; the supernatant was concentrated at 40°C by a rotary evaporator, and then, the concentrate was adjusted to pH 8.0 with Na_2HPO_4 (0.4 mol/L); second, the petroleum ether was added to the concentrate with the volume 1:1 (v/v) to decolor and repeated twice more; third, the residual petroleum ether in the concentrate was removed at 40°C by a rotary evaporator, and then, the concentrate was adjusted to pH 2.8 with citric acid (0.4 mol/L); fourth, the decolored concentrate was extracted thrice with ethyl acetate (1:1, v/v) and then concentrated at 40°C by a rotary evaporator; finally, the concentrated residue was increased to 5 mL with methanol and then kept at −20°C for quantification.

The extracts of endogenous hormones were first filtered with a durapore membrane (0.22 μm), and the samples (5 μL) were injected and quantified at 254 nm using an HPLC (Agilent 1260 Infinity II, CA, United States) with Symmetry C_{18} column (250 mm × 4.6 mm, 5.0 μm, CA, United States) at a column temperature of 25°C and flow rate of 1.0 mL/min. Methanol (A)–phosphoric acid (0.1% v/v, B) was the mobile phase with gradient elution: 0–1 min 10–45% (A), 1–3 min 45–55% (A), 3–5 min 55–65% (A), 5–8 min 65–75% (A), 8–12 min 75–10% (A), and 12–14 min 10–10% (A). The contents of GA_3, IAA, and ABA were quantified based on standard references (**Supplementary Figure 4**), and the standard curves were attached in **Supplementary Table 2**.

Statistical Analysis

All the measurements were performed using three biological replicates. Statistical analysis was performed *via* ANOVA and Duncan's multiple comparison tests, and SPSS 22.0 was the software package used with $P < 0.05$ as the basis for statistical differences. The Matrix Laboratory (MATLAB) software was used to program the linearization regression analysis for the relationship of EBF rate with vernalization temperature (VT), vernalization duration (VD), and seedlings classification (SC), with the number "1, 2, and 3" representing SC "small, medium, and large size," respectively.

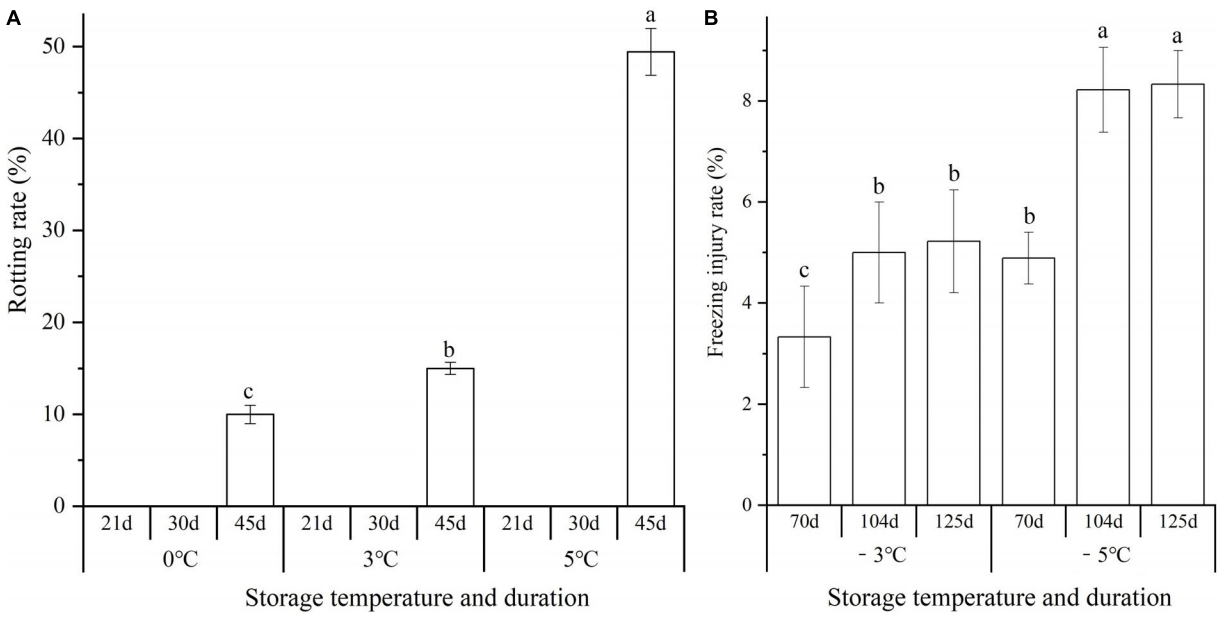


FIGURE 1 | Changes in the rotting rate (RR) and freezing injury rate (FIR) of seedlings at different storage temperatures and durations. Panels **(A,B)** represent the RR and FIR, respectively. Different lowercase letters represent a significant difference ($P < 0.05$) in the average value of different seedlings (large, medium, and small sizes) among different temperatures and durations.

RESULTS

Effect of Storage Temperature and Duration on Rotting and Freezing Injury

Significant differences in the rotting rate (RR) and freezing injury rate (FIR) of seedlings (including large, medium, and small sizes) were observed at different storage temperatures and durations. The rotting occurred after the seedlings were stored above 0°C for 45 d with the RR of 10.0, 15.0, and 49.4 at 0, 3, and 5°C, respectively (**Figure 1A**). The FIR at −3 and −5°C increased by 1.6- and 1.7-fold from 70 to 125 d, and the FIR at −5°C exhibited a 1.6, 3.2, and 3.1% higher than that of the −3°C at 70, 104, and 125 d, respectively (**Figure 1B**).

Effect of Storage Temperature and Duration on Germination

To investigate the effects of storage temperature and duration on the germination of seedlings, the germination rate (GR) was observed at 0, 3, 5, −3, and −5°C. There was a 1.2-, 1.2-, and 1.6-fold decrease in the average GR for the different seedlings

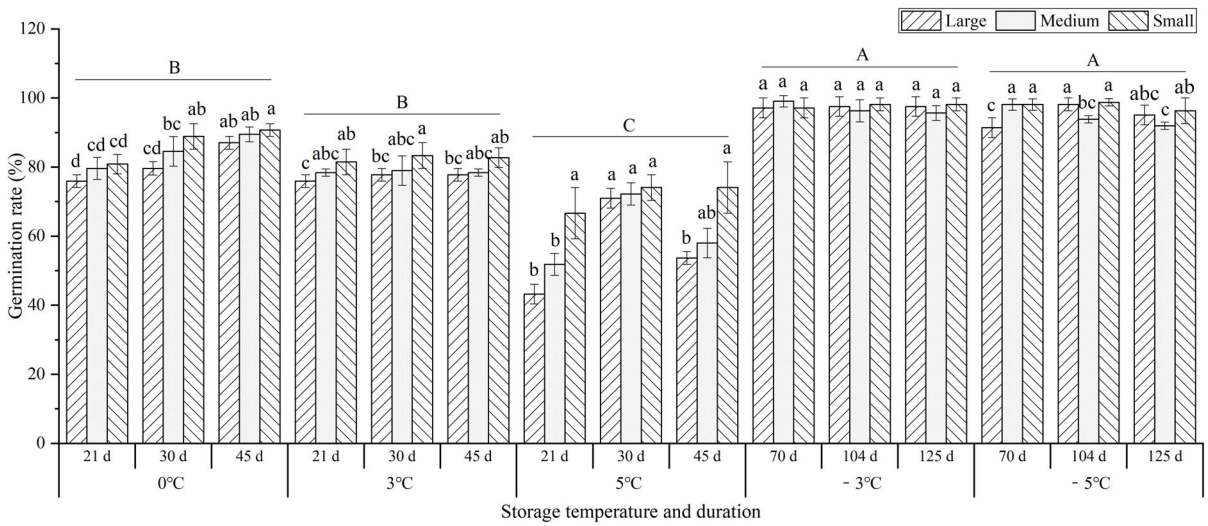


FIGURE 2 | Changes in the germination rate (GR) of seedlings at different storage temperatures and durations. Different uppercase letters represent a significant difference ($P < 0.05$) in the average value of different seedlings and storage durations among different temperatures. Different lowercase letters represent a significant difference between the size treatments at different temperatures.

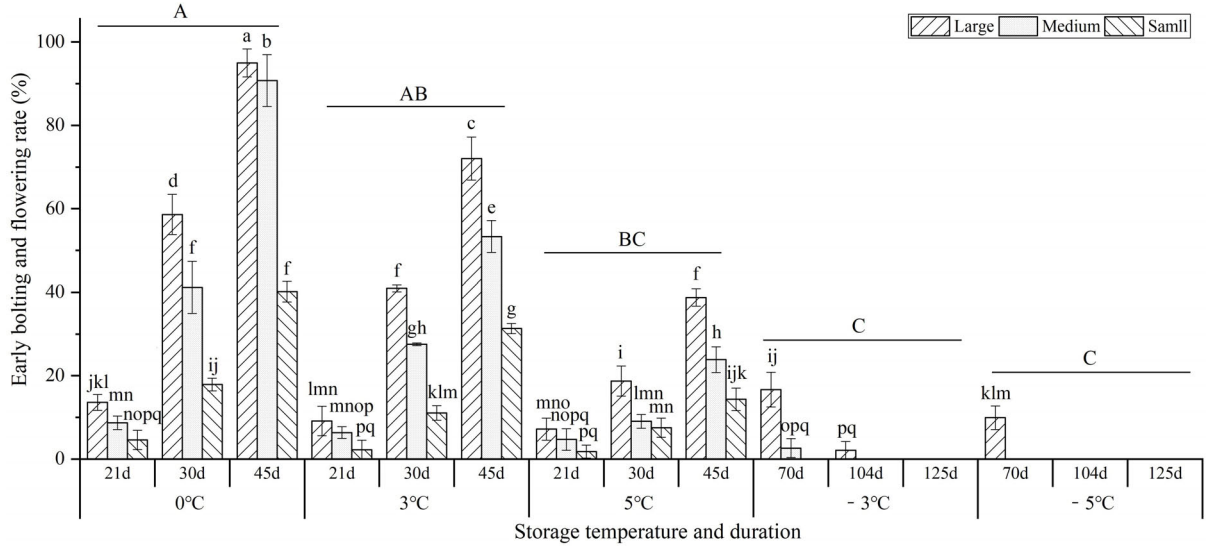


FIGURE 3 | Changes in the early bolting and flowering (EBF) rates of seedlings at different storage temperatures and durations. Different uppercase letters represent a significant difference ($P < 0.05$) in the average value of different seedlings and storage durations among different temperatures. Different lowercase letters represent a significant difference between the size treatments and storage durations among different temperatures.

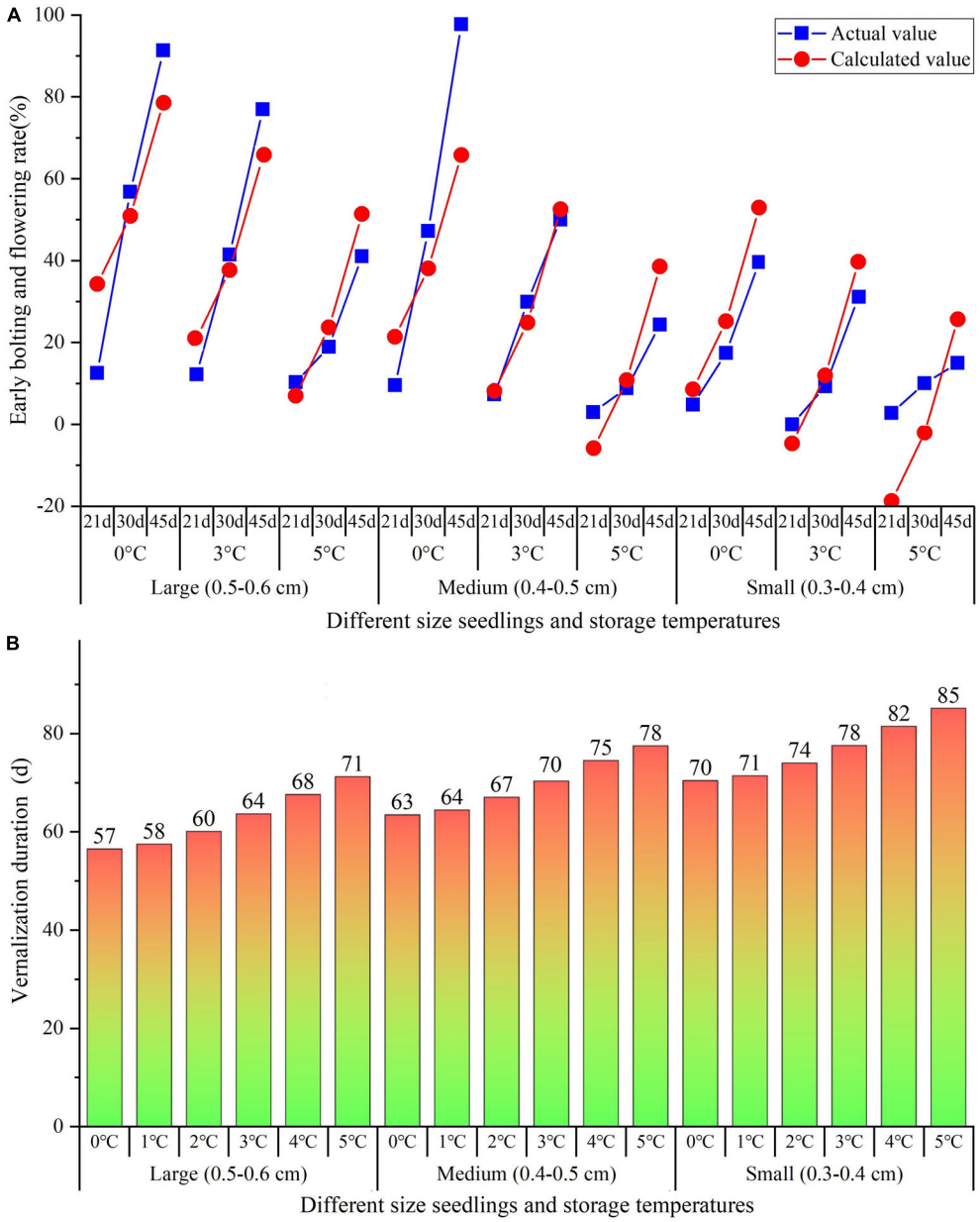


FIGURE 4 | Linearization regression analysis on the relationship of EBF rate with VT, VD, and SC **(A)**, as well as speculated the precise VT and VD of different size seedlings **(B)**. VT, vernalization temperature; VD, vernalization duration; and SC, seedlings classification.

(large, medium, and small sizes) at 0, 3, and 5°C compared to −3°C, respectively, but there was no significant difference between −3°C (97.4%) and −5°C (95.7%). In addition, the GR of the small size was higher than that of the large and medium sizes at 0, 3, and 5°C, but no obvious difference for large, medium, and small sizes was observed at −3 and −5°C (**Figure 2**).

Effect of Storage Temperature and Duration on Early Bolting and Flowering

As shown in **Figure 3**, storage temperature and duration had significant effects on the early bolting and flowering (EBF) rate. At the vernalization temperatures (0, 3, and 5°C), the EBF rates significantly increased with prolonged storage from 21 to 45 d, but significantly decreased with temperature increase; the highest EBF rate occurred at 0°C for 45 d with large, medium, and small sizes reaching 95.0, 90.8, and 40.2%, respectively. At the freezing temperatures (−3 and −5°C) for 70 d, the highest EBF rates for large sizes were 16.7 and 9.9%, but there was no EBF occurring for the medium and small sizes, which indicates that freezing storage (below −3°C) can effectively inhibit the EBF. In addition, the EBF rates were observed to be highest in the large size, followed by the medium and small sizes.

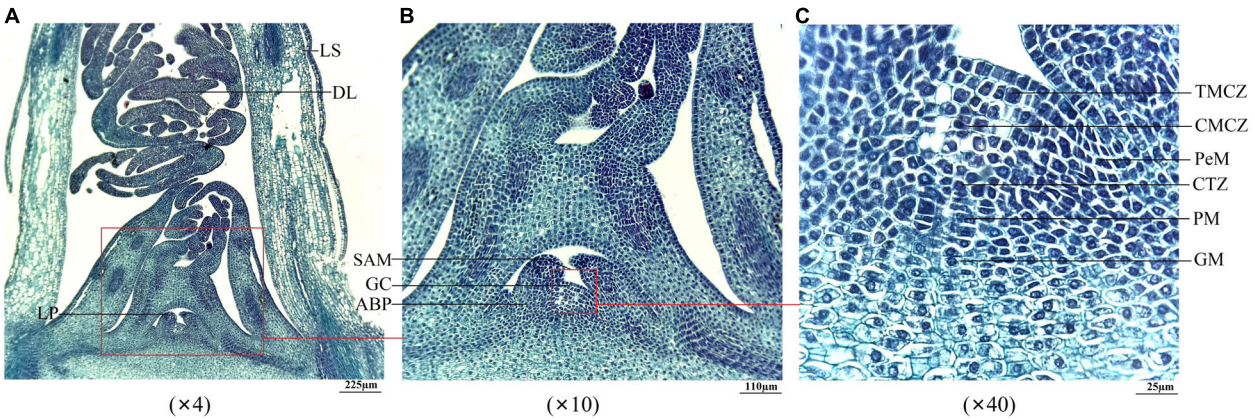


FIGURE 5 | Longitudinal anatomical structure of the SAM of seedlings of *Angelica sinensis*. Panels **(A–C)** represent the structure at 4×, 10×, and 40× magnification of the large size at the −5°C for 25 d, respectively. LS, leaf sheath; DL, delicate leaf; LP, leaf primordium; GC, growth cone; ABP, axillary bud primordium; TMCZ, tunica mother cell zone; CMCZ, central mother cell zone; PeM, peripheral meristem; CTZ, cambium-like transition zone; PM, pith meristem; GM, ground meristem.

Predicting the Precise Vernalization Temperature and Duration

Based on the linearization regression analysis on the relationship of EBF rate with VT, VD, and seedlings classification (SC), the regression model was obtained as the following equation:

$$EBF\ rate\ (\%) = 1.85 \times VD + 0.19 \times VT^3 - 2.04 \times VT^2 + 12.86 \times SC - 43.17$$

where *VD* (d), *VT* (0, 3, and 5°C), and *SC* (1, 2, and 3) with the number "1, 2, and 3" represent "small, medium, and large sizes," respectively.

In comparison of the actual EBF rate in this study with the calculated EBF rate by the equation, there was no difference between the actual and calculated curves that were close together (**Figure 4A**). Based on the above equation, the precise VD to complete the vernalization for flowering (EBF rate = 100%) was predicted with the VT ranging from 0 to 5°C and SC from 1 to 3 being available. As shown in **Figure 4B**, the VD ranges from 57 to 71 d with VT from 0 to 5°C for the large size, 63–78 d for the medium size, and 70–85 d for the small size.

Effect of Storage Temperature and Duration on Anatomical Structure of Shoot Apical Meristem

Since *A. sinensis* is a triennial plant in commercial cultivation, the vernalization of seedlings stored at the overwintering stage will confer to the ability to EBF in the second year (**Supplementary Figures 1–3**). To investigate the effects of storage temperature and duration on the SAM, first, the anatomical structures of the SAM were observed. As shown in **Figure 5**, the observed samples contained leaf sheath (LS), delicate leaves (DL), leaf primordium (LP), and SAM (**Figure 5A**); the SAM was composed of the growth cone (GC) and axillary bud primordium (ABP) (**Figure 5B**); the SAM that could be specifically distinguished at 40× magnification included tunica mother cell zone (TMCZ), central mother cell zone (CMCZ), peripheral meristem (PeM), cambium-like transition zone (CTZ), pith meristem (PM), and ground meristem (GM) (**Figure 5C**).

At 0, 3, and 5°C, obvious changes in the SAM were observed. At 0°C, the interval of a new LP formation was 20–30 d (i.e., 25, 50, 81, 104, and 125 d) for the large size, 30–35 d (i.e., 30, 60, 90, and 125 d) for medium size, and 45 d (i.e., 45 and 90 d) for small size (**Figure 6A**). At 3°C, the interval was 30 d (i.e., 30, 60, and 90 d) for the large size, 45 d (i.e., 45 and 90 d) for medium size, and 60 d (i.e., 30 and 90 d) for small size (**Figure 6B**). At 5°C, the interval was 30–45 d (i.e., 30, 45, and 75 d) for the large size, 45 d (i.e., 45 and 90 d) for medium size, and 76 d (i.e., 14 and 90 d) for small size (**Figure 6C**). At −3 and −5°C, little change was observed for the large, medium, and small sizes (**Supplementary Figure 6**). The specific anatomical structures of different size seedlings at different storage temperatures (0, 3, 5, −3, and −5°C) were shown in **Supplementary Figure 6**.

Effect of Storage Temperature and Duration on Root Activity

As shown in **Figure 7**, the root activity showed an increase trend with prolonged storage at 0 and 3°C, while it showed a decrease trend at the 5°C. At −3 and −5°C, the root activity maintained higher levels (over 105 $\mu g \cdot g^{-1} \cdot h^{-1}$) from 50 to 125 d than early stage 25 d (35 $\mu g \cdot g^{-1} \cdot h^{-1}$). Moreover, the root activity was significantly greater at the −3 and −5°C than that at 0, 3, and 5°C.

Effect of Storage Temperature and Duration on Primary Metabolism

For the soluble sugar contents, there was a 1.2-, 1.2-, and 1.1-fold decrease at 0, 3, and 5°C, respectively, from 14 to 45 d; there was a 1.5- and 1.5-fold decrease from 25 to 50 d, but a 1.8- and 1.9-fold increase from 50 to 125 d at the −3 and −5°C (**Figure 8A**). For the contents of starch and protein, decrease trends were observed at 0, 3, 5, −3, and −5°C, with 1.5-, 1.4-, 1.3-, 1.5-, and 1.2-fold

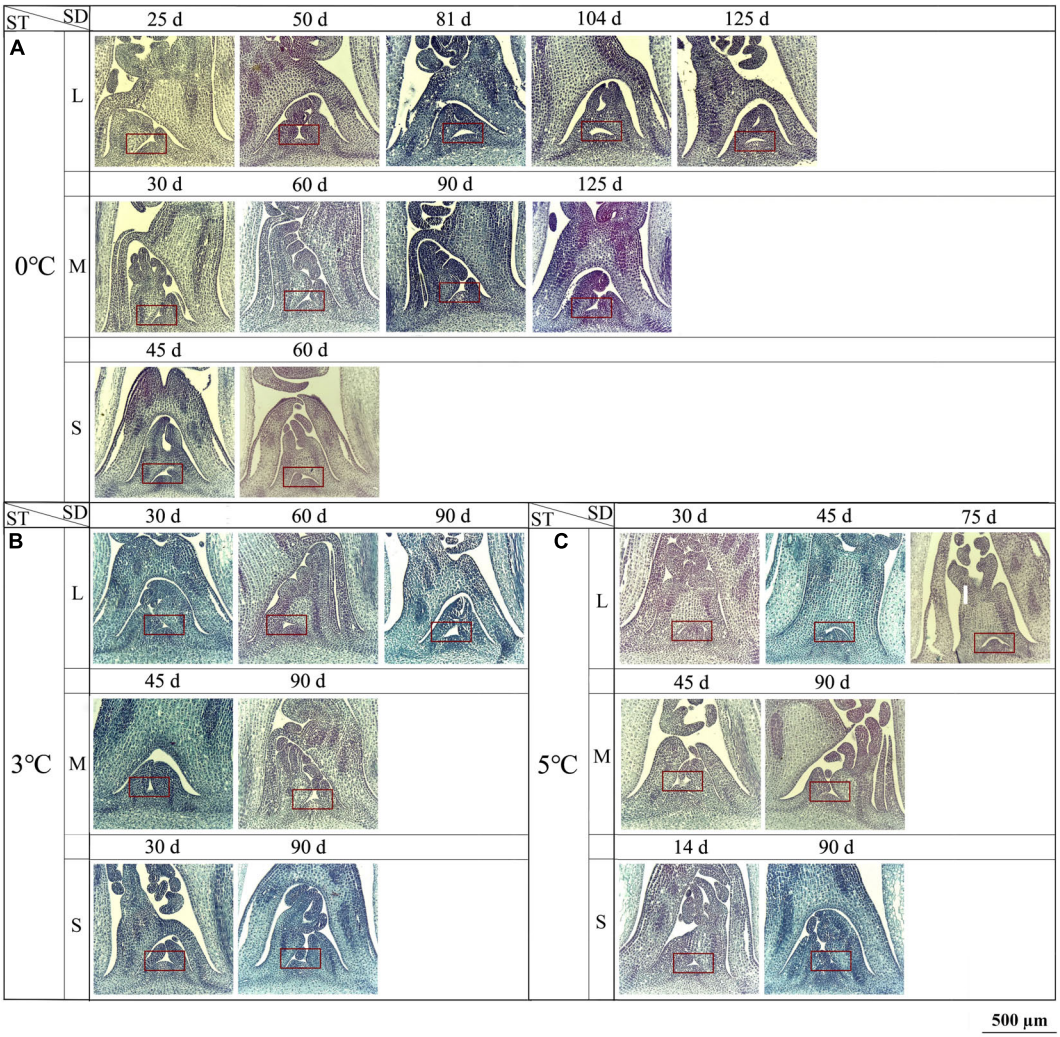


FIGURE 6 | Changes in longitudinal anatomical structure (10× magnification) of the SAM seedlings of *Angelica sinensis* at different storage temperatures and durations. The red frame shows the SAM. Panels **(A–C)** represent the structure of the SAM at 0, 3, and 5°C, respectively. SD, storage duration; ST, storage temperature; L, large size; M, medium size; S, small size. n = 10 per stage for the different size seedlings.

decrease for starch contents (**Figure 8B**) and a 1.1-, 1.2-, 1.3-, 1.1-, and 1.2-fold decrease for protein contents (**Figure 8C**). For the amino acid contents, a 1.2-, 1.6-, 1.4-, 1.4-, and 1.2-fold increase were observed at 0, 3, 5, −3, and −5°C, respectively (**Figure 8D**).

Effect of Storage Temperature and Duration on Endogenous Hormones

At 0, 3, and 5°C, the contents of GA_3 and IAA exhibited significant increases from 14 to 45 d, with 1.7-, 1.6-, and 1.6-fold increase for GA_3 (**Figure 9A**) and 1.4-, 1.8-, and 2.0-fold increase for IAA, respectively (**Figure 9B**), while at −3 and −5°C, the contents of GA_3 and IAA exhibited significant decreases from 25 to 125 d, with 1.7- and 1.7-fold decrease for GA_3 (**Figure 9A**) and 1.6- and 1.4-fold decrease for IAA, respectively (**Figure 9B**); for the ABA contents, significant decreases were observed at 0, 3, and 5°C, with 1.7-, 1.3-, and 1.6-fold decrease from 14 to 45 d; significant increases were observed at −3 and −5°C, with 1.4- and 1.4-fold increase from 25 to 125 d, respectively (**Figure 9C**).

DISCUSSION

Vernalization plays a crucial role in promoting plant flowering by prolonged exposure to low temperature (Amasino, 2005). Plants differ in the species and age at which they become sensitive to vernalization; thus, there are great differences in the cold requirement for flowering as well as anatomical structures and metabolites in the SAM; in addition, the vernalization can be lost at high temperature or avoided at freezing temperatures (Taiz and Zeiger, 2010b). For *A. sinensis*, the precise temperature and duration to complete the vernalization of seedlings have not been revealed, although several investigations strived to uncover the effective temperatures range for vernalization and find out the

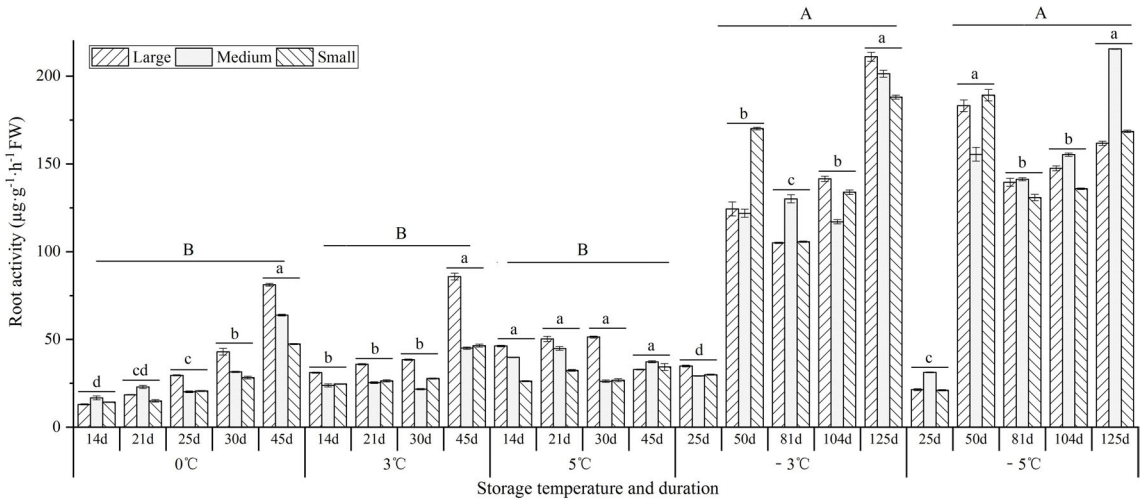


FIGURE 7 | Changes in the root activity of seedlings at different storage temperatures and durations. n = 10 per stage for the different size seedlings. Different uppercase letters represent a significant difference ($P < 0.05$) in the average value of different seedlings and storage durations among different temperatures. Different lowercase letters represent a significant difference in the average value of different seedlings between the storage durations for the same temperature.

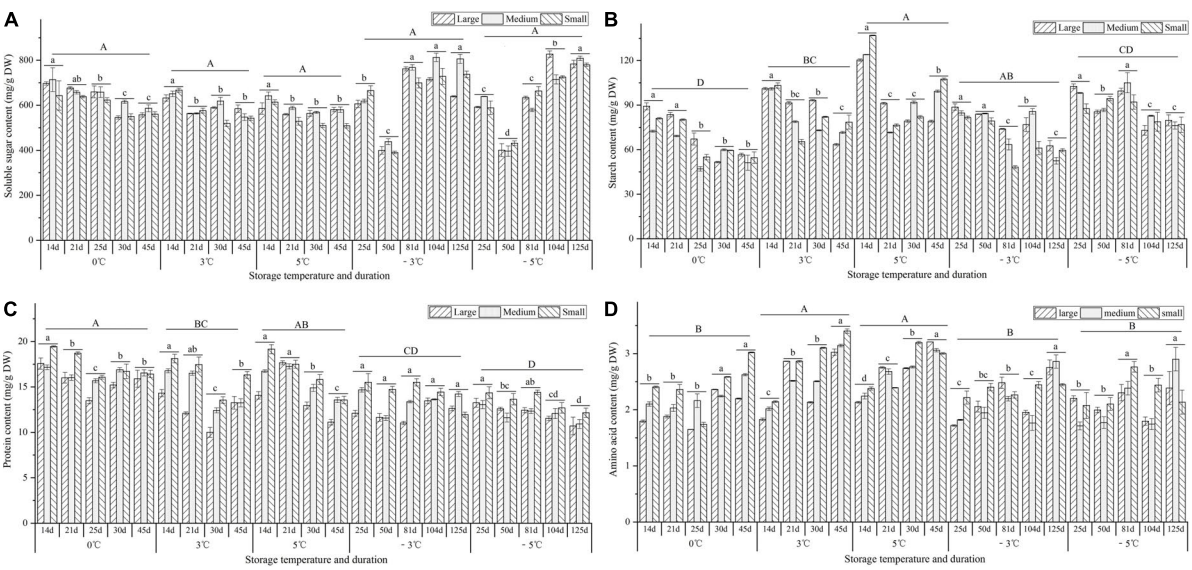


FIGURE 8 | Changes in the contents of soluble sugar, starch, protein, and amino acid of seedlings at different storage temperatures and durations. Panels **(A–D)** represent the contents of soluble sugar, starch, protein, and amino acid, respectively. n = 10 per stage for the different size seedlings. Different uppercase letters represent a significant difference ($P < 0.05$) in the average value of different seedlings and storage durations among different temperatures. Different lowercase letters represent a significant difference in the average value of different seedlings between the storage durations for the same temperature.

changes in metabolites (Wang, 1977; Li, 1979; Zhang and Huang, 1998; Chen et al., 2014; Jia et al., 2018). Here, we have found that vernalization temperature (VT) and duration (VD) are affected by the seedling sizes, the VD of seedlings ranges from 57 to 85 d with VT from 0 to 5°C, the early bolting and flowering (EBF) is effectively inhibited by seedlings were stored at the −3 to −5°C, and physiological characteristics significantly altered at different storages.

Extensive experiments have proved that temperature affects the quality of fruits, seeds, and seedlings during postharvest storage (Wilson et al., 2000; Cuadra-Crespo and Amor, 2010; Thomas et al., 2017). Rotting or chilling damages will occur when the tissues or organs are stored at non-freezing or freezing temperatures for a long time (Taiz and Zeiger, 2010a; Watkins, 2017). Previous studies have demonstrated that the seedlings of *A. sinensis* could germinate and grow after non-freezing (0 to 5°C) and freezing temperatures (−2 to −13°C) storage over 150 d (Wang, 1977, 1979; Wang and Zhang, 1982; Jia et al., 2018). In this study, the RR of seedlings increased with temperatures elevated from 0 to 5°C (*ca.* 49% at 5°C for 45 d), the FIR increased with temperatures declined from 0 to −5°C (*ca.* 8% at −5°C for 125 d) (see **Figure 1**), and the GR increased with temperatures

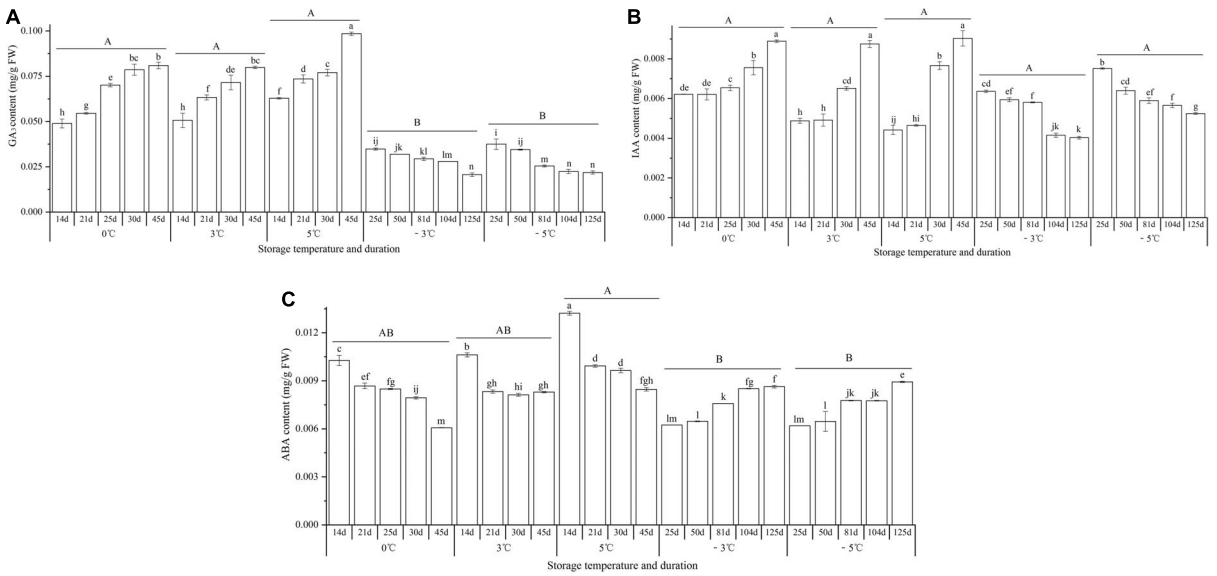


FIGURE 9 | Changes in the contents of GA_3, IAA, and ABA of seedlings at different storage temperatures and durations. Panels **(A–C)** represent the contents of GA_3, IAA, and ABA, respectively. $n = 10$ per stage for the different size seedlings. Different uppercase letters represent a significant difference ($P < 0.05$) in the average value of different storage durations for the medium seedlings among different temperatures. Different lowercase letters represent a significant difference between the storage durations for the same temperature.

declined from 5 to −5°C (*ca.* 94% at −5°C for 125 d) (see **Figure 2**).

The developmental signals that transit from vegetative growth to flowering include endogenous factors (e.g., circadian rhythms, phase change, and hormones) and external factors (e.g., photoperiodism and vernalization) (Taiz and Zeiger, 2010b). Furthermore, there is a linear relationship of apex diameter (AD) of *Brassica oleracea* with VT and VD, with the model $dAD/dVD = f[VT(VD)]$ (Wurr et al., 1993). For the *A. sinensis*, a positive relationship of the EBF rate with the size seedlings has been observed (Lin et al., 2007; Jia et al., 2018). The effective VT for seedlings ranged from 0 to 10°C, with an optimum range between 0 and 5°C (Wang, 1977, 1979; Jia et al., 2018). In this study, there was a linear relationship of EBF rate with VT, VD, and SC, and the precise VD was predicted to range from 57 to 85 d for different size seedlings at the VT from 0 to 5°C (see **Figure 4B**). Interestingly, the 52.2% EBF rate for the seedlings stored at 3°C for 45 d in this controlled experiment (see **Figure 3**) was almost consistent with *ca.* 40% EBF rate at 2.1°C for 33 d and 3.2°C for 44 d in the practical overwintered storage (see **Supplementary Figure 5** and **Supplementary Table 3**). In addition, several investigations have found that the EBF is also affected by the environmental factors (e.g., drought, latitude, and altitude), soil nutrients (e.g., nitrogen and phosphorus fertilizers), and planting density (Wang, 1977; Qi et al., 2004; Qiu et al., 2010).

Vernalization can be lost as a result of exposure to devernalizing conditions (e.g., high temperature) or freezing temperatures at which metabolic activity is suppressed (Taiz and Zeiger, 2010b). Previous studies on *A. sinensis* have found that the vernalization can be largely lost when seedlings exposure to 33–35°C for 2 d with the EBF rate of 20% or significantly avoided when seedlings exposure to −2 to −12°C for 210 d with the EBF rates 5.6–0% (Wang, 1979; Wang and Zhang, 1982). In this study, the vernalization can be avoided when the seedlings were stored at −3 to −5°C for 125 d, with no EBF occurrence (see **Figure 3**).

During the vernalization, stable changes (e.g., cell division, DNA replication, and gene expression) in the SAM are required (Taiz and Zeiger, 2010b; Dai et al., 2021). The SAM is a dynamic structure that changes during its cycle of leaf and stem formation, and the transition from vegetative to reproductive development is marked by an increase in the frequency of cell divisions within the central zone (Taiz and Zeiger, 2010b). Previous literature has reported that the number of *A. sinensis* leaves in one reproductive cycle always ranges from 12 to 14, and a physiological age developing into flowering would be advanced with one more new leaf producing (Tang, 1980). In this study, there were regular structural changes in the SAM for the different seedlings exposure to different temperatures and durations, with more leaves producing at the VT (0, 3, and 5°C), while little change at freezing temperatures (−3 and −5°C) (see **Figure 6**), which shows that the vernalization accelerates the leaves formation and eventually the physiological age. Similar results of morphological changes in the SAM have also been observed in other plants in response to vernalization, such as *Lolium temulentum*, *Antirrhinum majus*, *Brunonia australis,* and *Calandrinia* (Arumuganathan et al., 1991; Adams et al., 2003; Cave et al., 2011). The acceleration of growth and development can be induced from the differential expression of genes, such as *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*), *AGAMOUS-LIKE19* (*AGL19*), and *FLOWERING LOCUS C* (*FLC*) (Sheldon et al., 2000; Taiz and Zeiger, 2010b; Suter et al., 2014; Jaudal et al., 2018).

During the vernalization, active metabolism in the SAM is also required (Taiz and Zeiger, 2010b). Previous studies on *A. sinensis* found that the soluble sugar content decreased, but the nitrate reductase activity increased after the seedlings exposure to 4°C for 170 d (Zhang and Huang, 1998); the contents of soluble sugar and protein decreased, but the malonic dialdehyde content increased after the seedlings exposure to −10°C for 210 d (Chen et al., 2014). In this study, the levels of soluble sugar, starch, protein, and ABA in the seedlings decreased, but the levels of root activity, amino acid, GA_3, and IAA exhibited an increase trend after the seedlings were stored at the VT (0, 3, and 5°C) for 45 d; the levels of root activity, soluble sugar, amino acid, and ABA increased, but the levels of starch, protein, GA_3, and IAA decreased after the seedlings were stored at freezing temperatures (−3 and −5°C) for 125 d (see **Figures 7**–**9**). These changes in physiological characteristics can well explain the previously measured index, for example, the root activity is in accordance with the germination rate, the levels of GA_3 and IAA involved in growth and development are in accordance with the EBF rate, and the levels of soluble sugar, amino acid, and ABA involved in stress tolerance are in accordance with the freezing injury rate. Similar results of physiological changes in the seedlings have also been observed in other plants in response to vernalization and freezing temperatures, such as a decrease in starch and total nitrogen in cabbage after vernalization (Zhao et al., 2010), a general increase in neutral and acidic amino acids in both spring and winter wheat varieties grown at 2°C for 2 weeks (Trione et al., 1967), a significant increase in GA and related biosynthetic genes (*ent-kaurene oxidase* and *GA20-oxidase*) in Pak Choi after 4°C treatment (Shang et al., 2017), an increase in IAA and related biosynthetic genes (*IAA8* and *nitrilase/nitrile hydratase NIT4*), and a decrease in ABA and related biosynthetic genes (*ABA2* and *9-cis-epoxycarotenoid dioxygenase NCED1*) in *Beta vulgaris* after 4°C treatment for 16 weeks (Liang et al., 2018).

CONCLUSION

From the above observations, the vernalization of *A. sinensis* seedlings is dynamically regulated by temperature and duration, and the EBF rate can be effectively inhibited with the seedlings stored below freezing temperatures. Based on the linearization regression analysis, a model of the EBF rate with vernalization temperature and vernalization duration has been proposed for different size seedlings. The epigenetic regulation mechanism of vernalization will be examined *via* future transcriptomics and epigenomics investigations.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XL and MML: investigation, data curation, and writing—original draft. MFL: conceptualization, writing—review and editing, and project administration. JW: conceptualization and project administration. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by funds from the State Key Laboratory of Aridland Crop Science Gansu Agricultural University (GSCS-2021-Z03), National Natural Science Foundation of China (32160083), China Agriculture Research System of MOF and MARA (CARS-21), and Key talent projects of Gansu Province (2020RCXM103).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.853444/full#supplementary-material

REFERENCES

Adams, S. R., Munir, M., Valdés, V. M., Langton, F. A., and Jackson, S. D. (2003). Using flowering times and leaf numbers to model the phases of photoperiod sensitivity in *Antirrhinum majus* L. *Ann. Bot.* 92, 689–696. doi: 10.1093/aob/mcg194

Amasino, R. M. (2005). Vernalization and flowering time. *Curr. Opin. Biotech.* 16, 154–158. doi: 10.1016/j.copbio.2005.02.004

Arumuganathan, K., Dale, P. J., and Cooper, J. P. (1991). Vernalization in *Lolium temulentum* L.: responses of in vitro cultures of mature and immature embryos shoot apices and callus. *Ann. Bot.* 67, 173–179. doi: 10.1093/oxfordjournals.aob.a088117

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. *Anal. Chem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3

Cao, X. C., Zhu, C. Q., Zhong, C., Hussain, S., Zhu, L. F., Wu, L. H., et al. (2018). Mixed-nitrogen nutrition-mediated enhancement of drought tolerance of rice seedlings associated with photosynthesis, hormone balance and carbohydrate partitioning. *Plant Growth Regul.* 84, 451–465. doi: 10.1007/s10725-017-0352-6

Cave, R. L., Birch, C. J., Hammer, G. L., Erwin, J. E., and Johnston, M. E. (2011). Juvenility and flowering of *Brunonia australis* (Goodeniaceae) and *Calandrinia* sp. (Portulacaceae) in relation to vernalization and daylength. *Ann. Bot.* 108, 215–220. doi: 10.1093/aob/mcr116

Chao, W. W., Hong, Y. H., Chen, M. L., and Lin, B. F. (2010). Inhibitory effects of *Angelica sinensis* ethyl acetate extract and major compounds on NF-κB trans-activation activity and LPS-induced inflammation. *J. Ethnopharmacol.* 129, 244–249. doi: 10.1016/j.jep.2010.03.022

Chen, H. G., Du, T., Zhu, T. T., Gao, S. F., Chai, L., and He, W. W. (2014). Study on physiological mechanisms in frozen storage to reduce early bolting of *Angelica sinensis*. *Mod. Tradit. Chin. Med. Mater. Med. World Sci. Tech.* 16, 203–206. doi: 10.11842/wst.2014.01.038

Committee for the Pharmacopoeia of PR China (2015). *Pharmacopoeia Of The People's Republic of China*, Vol. 1. Beijing: Chinese Medical Science and Technology Press, 133.

Cuadra-Crespo, P., and Amor, F. M. D. (2010). Effects of postharvest treatments on fruit quality of sweet pepper at low temperature. *J. Sci. Food Agr.* 90, 2716–2722. doi: 10.1002/jsfa.4147

Dai, Y., Sun, X., Wang, C. G., Li, F., Zhang, S. F., Zhang, H., et al. (2021). Gene co-expression network analysis reveals key pathways and hub genes in Chinese cabbage (*Brassica rapa* L.) during vernalization. *BMC Genomics* 22:236. doi: 10.1186/s12864-021-07510-8

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., and Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28, 350–356. doi: 10.1021/ac60111a017

Huang, L. Q., and Jin, L. (2018). *Suitable Technology For Production And Processing Of Angelica Sinensis*. Beijing: China Pharmaceutical Science and Technology Press, 1–14.

Jaudal, M., Zhang, L., Che, C., Li, G., Tang, Y., Wen, J., et al. (2018). A *SOC1-like* gene *MtSOC1a* promotes flowering and primary stem elongation in Medicago. *J. Exp. Bot.* 69, 4867–4880.

Jia, Z., Di, S. Q., Zhang, J. J., Zhao, F. Y., and Li, S. X. (2017). Effect of overwintering low temperatures and planting dates on *Angelica sinensis* growth cultivated at lower altitude region. *J. Northwest A F Univ. (Nat. Sci. Ed.)* 45, 129–135, 144. doi: 10.1093/jxb/ery284

Jia, Z., Di, S. Q., Zhao, F. Y., Li, S. X., and Wang, S. C. (2018). Effects of different low overwintering temperatures on *Angelica* vernalization and premature bolting. *Agr. Sci. Tech.* 19, 55–62. doi: 10.16175/j.cnki.1009-4229.2018.04.008

Li, G. J. (1979). Study on biological characteristics of *Angelica sinensis* and the prevention of early bolting. *J. Chin. Med. Mat.* 2, 1–10. doi: 10.13863/j.issn1001-4454.1979.02.001

Li, J., Li, M. L., Zhu, T. T., Zhang, X. N., Li, M. F., and Wei, J. H. (2021). Integrated transcriptomics and metabolites at different growth stages reveals the regulation mechanism of bolting and flowering of *Angelica sinensis*. *Plant Biol.* 23, 574–582. doi: 10.1111/plb.13249

Li, M. F., Kang, T. L., Jin, L., and Wei, J. H. (2020a). Research progress on bolting and flowering of *Angelica sinensis* and regulation pathways. *Chin. Tradit. Herbal Drugs* 51, 5894–5899. doi: 10.7501/j.issn.0253-2670.2020.22.029

Li, M. F., Liu, X. Z., Wei, J. H., Zhang, Z., Chen, S. J., Liu, Z. H., et al. (2020b). Selection of high altitude planting area of Angelica sinensis based on biomass, bioactive compounds accumulation and antioxidant capacity. *Chin. Tradit. Herbal Drugs* 51, 474–481. doi: 10.7501/j.issn.0253-2670.2020.02.026

Li, M. F., Lv, M., Yang, D. L., Wei, J. H., Xing, H., and Pare, P. W. (2020c). Temperature-regulated anatomical and gene-expression changes in sinopodophyllum hexandrum seedlings. *Ind. Crop. Prod.* 152, 112479. doi: 10.1016/j.indcrop.2020.112479

Li, M. L., Cui, X. W., Jin, L., Li, M. F., and Wei, J. H. (2022). Bolting reduces ferulic acid and flavonoid biosynthesis and induces root lignification in *Angelica sinensis*. *Plant Physiol. Biochem.* 170, 171–179. doi: 10.1016/j.plaphy.2021.12.005

Li, X. F., and Zhang, Z. L. (2016). *Guidance for Plant Physiology Experiments*, 5th Edn. Beijing: Higher Education Press, 30–31.

Li, Y., Huang, H. P., Deng, R., Zhang, X. C., Cui, Y. M., Lu, X. L., et al. (2016). The improvement of the sample preparation technique for TEM of plant. *Chin. J. Trop. Crops* 37, 2100–2105. doi: 10.3969/j.issn.1000-2561.2016.11.011

Liang, N. G., Cheng, D. Y., Liu, Q. H., Luo, C. F., and Dai, C. H. (2018). Vernalization and photoperiods mediated IAA and ABA synthesis genes expression in *Beta vulgaris*. *Russ. J. Plant. Physiol.* 65, 642–650. doi: 10.1134/S1021443718050126

Lin, H. M., Qiu, D. Y., and Chen, Y. (2007). Effect of root diameter on early bolting rate and yield in seedling of *Angelica sinensis*. *Chin. Tradit. Herbal Drugs* 38, 1386–1389. doi: 10.3321/j.issn:0253-2670.2007.09.038

Ma, J. P., Guo, Z. B., Jin, L., and Li, Y. D. (2015). Phytochemical progress made in investigations of *Angelica sinensis* (Oliv.) Diels. *Chin. J. Nat. Med.* 13, 241–249. doi: 10.1016/S1875-5364(15)30010-8

McGuffin, M., Kartesz, J. T., Leung, A. Y., and Tucker, A. O. (2000). *Herbs of Commerce*, 2nd Edn. Silver Spring, MD: American Herbal Products Association, 421.

Pan, X. Q., Welti, R., and Wang, X. M. (2010). Quantitative analysis of major plant hormones in crude plant extracts by high-performance liquid chromatography-mass spectrometry. *Nat. Protoc.* 5, 986–992. doi: 10.1038/nprot.2010.37

Qi, J. T., Lin, H. M., and Liu, X. Z. (2004). Preliminary study on the effect of nitrogen and phosphorus fertilizer on bolting rate of *Angelica sinensis*. *J. Chin. Med. Mat.* 27, 82–83. doi: 10.3321/j.issn:1001-4454.2004.02.003

Qiu, D. Y., Lin, H. M., Chen, Y., Li, Y. D., and Guo, F. X. (2010). Effects of latitude, longitude and altitude on *Angelica* growth and early bolting in medicine formation period. *Acta Agrestia Sin.* 18, 838–843. doi: 10.3724/SP.J.1077.2010.01263

Shang, M. Y., Wang, X. T., Zhang, J., Qi, X. H., Ping, A. M., Hou, L. P., et al. (2017). Genetic regulation of GA metabolism during vernalization, floral bud initiation and development in Pak Choi (*Brassica rapa* ssp. chinensis Makino). *Front. Plant Sci.* 8:1533. doi: 10.3389/fpls.2017.01533

Sheldon, C. C., Rouse, D. T., Finnegan, E. J., Peacock, W. J., and Dennis, E. S. (2000). The molecular basis of vernalization: The central role of flowering locus C (FLC). *Proc. Natl. Acad. Sci. U.S.A.* 97, 3753–3758. doi: 10.1073/pnas.97.7.3753

Suter, L., Rüegg, M., Zemp, N., Hennig, L., and Widmer, A. (2014). Gene regulatory variation mediates flowering responses to vernalization along an altitudinal gradient in *Arabidopsis*. *Plant Physiol.* 166, 1928–1942. doi: 10.1104/pp.114.247346

Taiz, L., and Zeiger, E. (2010b). "Plant physiology," in *The Control Of Flowering*, 5th Edn, eds D. E. Fosket and R. Amasino (Sunderland, MA: Sinauer Associates, Inc), 559–590.

Taiz, L., and Zeiger, E. (2010a). "Plant physiology," in *Responses and Adaptation To Abiotic Stress*, 5th Edn, eds R. Bressan, M. C. Drew, P. M. Hasegawa, R. Locy, V. M. Michael, and D. E. Salt (Sunderland, MA: Sinauer Associates, Inc), 609–610.

Tang, X. Y. (1980). Study on controlling early bolting of *Angelica sinensis* (Oliv.). *Diels. J. Mod. Clin. Med.* 6, 12–13.

Thomas, B., Murray, B. G., and Murphy, D. J. (2017). "Encyclopedia of applied plant sciences," in *Postharvest Physiological Disorders of Fresh Crops*, 2nd Edn. ed. L. Colville (Amsterdam: Elsevier), 335–339.

Trione, E. J., Young, J. L., and Yamamoto, M. (1967). Free amino acid changes associated with vernalization of wheat. *Phytochemistry* 6, 85–91. doi: 10.1016/0031-9422(67)85011-8

Upton, R. (2003). *American Herbal Pharmacopoeia And Therapeutic Compendium: Dang Gui Root-Angelica Sinensis (Oliv.)*. Scotts Valley, CA: American Herbal Pharmacopoeia, 1–41.

Wang, B. H., and Ou-Yang, J. P. (2005). Pharmacological actions of sodium ferulate in cardiovascular system. *Cardiovasc. Drug Rev.* 23, 161–172. doi: 10.1111/j.1527-3466.2005.tb00163.x

Wang, L. Y., Tang, Y. P., Liu, X., Zhu, M., Tao, W. W., Li, W. X., et al. (2015). Effects of ferulic acid on antioxidant activity in *Angelicae sinensis* Radix, Chuanxiong Rhizoma, and their combination. *Chin. J. Nat. Medicines* 13, 401–408. doi: 10.1016/S1875-5364(15)30032-7

Wang, W. J. (1977). Analysis and control of early bolting characteristic of *Angelica sinensis*. *J. Northwest Univ. (Nat. Sci. Ed.)* 7, 32–39.

Wang, W. J. (1979). Technology and principle of seedling frozen storage of *Angelica sinensis*. *J. Chin. Med. Mat.* 3, 1–5.

Wang, W. J., and Zhang, Z. M. (1982). Bolting characteristics and control approaches of *Angelica sinensis*. *J. Northwest Plant Res.* 2, 95–104.

Wang, X. K. (2006). *Experimental Guidance of Plant Physiology Module*. Beijing: Science Press, 174–241.

Watkins, C. B. (2017). "Postharvest physiological disorders of fresh crops-sciencedirect," in *Encyclopedia of Applied Plant Sciences*, 2nd Edn, ed. B. Thomas (Amsterdam: Elsevier), 315–322. doi: 10.1016/b978-0-12-394807-6.00217-3

Wei, W. L., Zeng, R., Gu, C. M., Qu, Y., and Huang, L. F. (2016). *Angelica sinensis* in China-A review of botanical profile, ethnopharmacology, phytochemistry and chemical analysis. *J. Ethnopharmacol.* 190, 116–141. doi: 10.1016/j.jep.2016.05.023

Wilson, S. B., Rajapakse, N. C., and Young, R. E. (2000). Use of low temperature to improve storage of *in vitro* Broccoli seedlings under various light qualities. *J. Veg. Crop Prod.* 6, 51–67. doi: 10.1300/j068v06n02_07

Wurr, D. C. E., Fellows, J. R., Phelps, K., and Reader, R. J. (1993). Vernalization in summer/autumn cauliflower (*Brassica oleracea* var. *botrytis* L.). *J. Exp. Bot.* 44, 1507–1514. doi: 10.1093/jxb/44.9.1507

Xu, X. Q., Zhang, X. B., Chen, J., Zhao, W. L., and Jin, L. (2020). Study on ecological suitability of *Angelica sinensis* in Gansu Province. *Chin. Tradit. Herbal Drugs* 51, 3304–3307. doi: 10.7501/j.issn.0253-2670.2020.12.024

Yang, D. L., Sun, P., and Li, M. F. (2016). Chilling temperature stimulates growth, gene overexpression and podophyllotoxin biosynthesis in *Podophyllum hexandrum* Royle. *Plant Physiol. Biochem.* 107, 197–203. doi: 10.1016/j.plaphy.2016.06.010

Yao, L. (2005). Effect of shading during the nursery of *Angelica sinensis* on bolting rate and economic characters. *Gansu Agr. Sci. Tech.* 10, 54–55. doi: 10.3969/j.issn.1001-1463.2005.10.027

Zhang, E. H., and Huang, P. (1998). Effects of vernalization treatment on physiological character of *Angelica sinensis* seedlings. *J. Gansu Agr. Univ.* 33, 240–243.

Zhang, H. Y., Bi, W. G., Yu, Y., and Liao, W. B. (2012). *Angelica sinensis* (Oliv.) diels in china: distribution, cultivation, utilization and variation. *Genet. Resour. Crop Evol.* 59, 607–613. doi: 10.1007/s10722-012-9795-9

Zhang, S. Y., and Cheng, K. C. (1989). "*Angelica sinensis* (Oliv.) diels: *in vitro* culture, regeneration, and the production of medicinal compounds," in *Medicinal And Aromatic Plants II. Biotechnology In Agriculture and Forestry*, Vol. 7, ed. Y. P. S. Bajaj (Berlin: Springer), 1–22.

Zhang, Z., Liu, X. Z., Bao, Y. J., Li, J., Cao, X. L., and Li, M. F. (2018). Selection of cropping rotations of *Angelica sinensis* based on biomass, bioactive compounds accumulation and antioxidant capacity. *J. Gansu Agr. Univ.* 53, 86–89.

Zhao, R. Q., Jiang, X. M., and Yu, X. H. (2010). Effect of low temperature vernalization on metabolism of carbon and nitrogen of cabbage. *J. Northeast Agr. Univ.* 41, 38–42.

Conflict of Interest: 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.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 23 May 2022
doi: 10.3389/fpls.2022.882965

Low Female Gametophyte Fertility Contributes to the Low Seed Formation of the Diploid Loquat [*Eriobotrya Japonica (Thunb.)* Lindl.] Line H30-6

Qingqing Xia[1,2†], Jiangbo Dang[1,2†], Peng Wang[1,2], Senlin Liang[3], Xu Wei[4], Xiaolin Li[1,2], Suqiong Xiang[1,2], Haiyan Sun[1,2], Di Wu[1,2], Danlong Jing[1,2], Shumin Wang[1,2], Yan Xia[1,2], Qiao He[1,2], Qigao Guo[1,2] and Guolu Liang[1,2*]*

[1] Key Laboratory of Horticulture Science for Southern Mountains Regions of Ministry of Education, College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China, [2] State Cultivation Base of Crop Stress Biology for Southern Mountainous Land of Southwest University, Academy of Agricultural Sciences, Southwest University, Chongqing, China, [3] Economic Crops of Ziyang City, Ziyang City, China, [4] America Citrus Research and Education Center, University of Florida, Gainesville, FL, United States

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Feng Ding,
Guangxi Academy of Agricultural Science, China
Aytekin Polat,
Mustafa Kemal University, Turkey

***Correspondence:**
Qigao Guo
qgguo@126.com
Guolu Liang
lianggl@swu.edu.cn

[†] These authors have contributed equally to this work and share first authorship

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

Received: 24 February 2022
Accepted: 19 April 2022
Published: 23 May 2022

Citation:
Xia Q, Dang J, Wang P, Liang S, Wei X, Li X, Xiang S, Sun H, Wu D, Jing D, Wang S, Xia Y, He Q, Guo Q and Liang G (2022) Low Female Gametophyte Fertility Contributes to the Low Seed Formation of the Diploid Loquat [Eriobotrya Japonica (Thunb.) Lindl.] Line H30-6. Front. Plant Sci. 13:882965. doi: 10.3389/fpls.2022.882965

Loquat is a widely grown subtropic fruit because of its unique ripening season, nutrient content, and smooth texture of its fruits. However, loquat is not well-received because the fruits contain many large seeds. Therefore, the development of seedless or few-seed loquat varieties is the main objective of loquat breeding. Polyploidization is an effective approach for few-seed loquat breeding, but the resource is rare. The few-seed loquat line H30-6 was derived from a seedy variety. Additionally, H30-6 was systematically studied for its fruit characteristics, gamete fertility, pollen mother cell (PMC) meiosis, stigma receptivity, *in situ* pollen germination, fruit set, and karyotype. The results were as follows. (1) H30-6 produced only 1.54 seeds per fruit and the fruit edible rate was 70.77%. The fruit setting rate was 14.44% under open pollination, and the other qualities were equivalent to those of two other seedy varieties. (2) The *in vitro* pollen germination rate was only 4.04 and 77.46% of the H30-6 embryo sacs were abnormal. Stigma receptivity and self-compatibility in H30-6 were verified by *in situ* pollen germination and artificial pollination. Furthermore, the seed numbers in the fruits of H30-6 did not significantly differ among any of the pollination treatments (from 1.59 ±0.14 to 2 ± 0.17). (3) The chromosome configuration at meiotic diakinesis of H30-6 was 6.87I + 9.99II + 1.07III +0.69IV +0.24V (H30-6), and a total of 89.55% of H30-6 PMCs presented univalent chromosomes. Furthermore, chromosome lagging was the main abnormal phenomenon. Karyotype analysis showed that chromosomes of H30-6 had no recognizable karyotype abnormalities leading to unusual synapsis on the large scale above. (4) The abnormal embryo sacs of H30-6 could be divided into three main types: those remaining in the tetrad stage (13.38%), those remaining in the binucleate embryo sac stage (1.41%), and those without embryo sacs (52.82%). Therefore, we conclude that the loquat line H30-6 is a potential few-seed loquat resource. The diploid loquat

line H30-6 was with low gametophyte fertility, which may be driven by abnormal meiotic synapses. The low female gamete fertility was the main reason for the few seeds. This diploid loquat line provides a new possibility for breeding a few-seed loquat at the diploid level.

Keywords: seedless fruit, male sterility, female sterility, gametophyte differentiation and development, meiosis

INTRODUCTION

Loquat (*Eriobotrya japonica*), which originated in China, is a globally grown subtropical fruit and its fruits have high nutrition and a delicious taste, consumed directly or made into juices, jams, wines, and syrups (Lin et al., 1999; Lin, 2008). There are large seeds in loquat fruits, which is inconvenient for eating and processing. Therefore, the development of seedless and few-seed varieties has been the main goal of loquat breeders.

For a long period of time, polyploidization has been an effective approach to breeding seedless and few-seed loquat varieties. Early in 1984, Huang et al. (1984) reported that a tetraploid line ('Min No. 3') was obtained *via* colchicine induction, and the majority of fruits (91.3%) had only one seed. In the 1990s, triploids were employed to produce seedless loquats by Liang, and seedless varieties including "Wuhe Guoyu" (Guo et al., 2016), "Huajinwuhe No. 1" (Dang et al., 2019a), and "Huayu Wuhe No. 1" (Dang et al., 2019b) have since been released. However, hormone treatment is necessary to increase the fruit setting rate of triploid loquat varieties (Maria et al., 2013). Some triploid lines were found to produce few-seed fruits when pollinated by diploid lines without hormone treatment (Kikuchi et al., 2014; Yang, 2021). In recent years, some tetraploids have been applied to breeding triploids on a large scale (Liang et al., 2018; Wang et al., 2021).

At the diploid level, seedless fruits could be obtained *via* hormone regulation, but the fruit qualities have been proven to be significantly lower than those of the seeded fruits (Gu and Zhang, 1990; Zhang et al., 1999; Chen et al., 2006; Deng, 2009b; Maria et al., 2013). "Tai Cheng No. 4" was reported to be a loquat variety with 1.3 seeds per fruit, and fruits with one seed were observed for up to 72.6% (Huang and Xu, 1980). However, there is no further information about this variety. An investigation reported by Jiang et al. (2009) indicated that 15 out of 128 lines bear fruits with 1–2 seeds, and there are two of the lowest lines (Guangmian and Hongxing), with 1.2 seeds per fruit. Zhang et al. (2014) reported that there were only 1.12 seeds per fruit of "Chuannong No.1"; however, the fruit setting rate of this variety was very low. Therefore, it can be concluded based on the above information that few-seed loquat resources are scarce.

It is logical that the abnormal meiosis of polyploids contributes to low fertility gametes, thus leading to seedless or few-seed fruit (Maria et al., 2013; Xu et al., 2016). But the few-seed mechanism of diploid loquat may be a more complex issue. While there have only been a few reports concerning this issue. Zhang et al. (2014, 2015) reported that gametophytic self-incompatibility resulted in the few-seed of "Chuannong No. 1", and pollen inactivation caused by high temperatures contributed to a low fruit setting rate and few-seed fruit of "Chuanzao". Mutant shoot breeding, chance seedling selection, and cross-breeding play an important role in the seedless variety breeding (Ledbetter and Ramming, 1989; Deng et al., 1996; Fan et al., 2004; Xu et al., 2006; Ye et al., 2006; Dong et al., 2013; Li and Wang, 2019; Li et al., 2019). So, it is important to discover more few-seed germplasm and identify the mechanism underlying the production of low numbers of seeds for breeding few-seed loquats.

An open-pollinated progeny of a white-fleshed variety "Huabai No.1", named H30-6, was shown to bear fruits with few seeds. This line thus is an ideal material to breed new few-seed varieties and to research the mechanism underlying low seed formation. In the current study, we observed the stability of the H30-6 fruit set and quality traits over successive years. Moreover, the flower structure, gamete fertility, pollen mother cell (PMC) meiosis, stigma receptivity, *in situ* pollen germination, fruit setting, karyotypes, and so forth were studied to determine the factors underlying the low seed formation of loquat line H30-6. Taken together, the results can be referred to in few-seed loquat variety breeding at the diploid level.

MATERIALS AND METHODS

Plant Materials

From 2016 to 2021, we obtained materials from vigorously growing plants in the orchard of fruit germplasm resources belonging to the Key Fruit Lab, of the College of Horticulture and Landscape Architecture, Southwest University. H30-6 is an open-pollinated progeny of a white-fleshed variety "Huabai No.1", which is red-fleshed. In the orchard, the H30-6 line was represented by one 10-year-old individual and five 3-year-old grafted trees. So, in this study, in addition to the H30-6 line (red-fleshed), we chose two varieties differing in fertility, namely, "Huabai No. 1" (white-fleshed, named H411) and "Jinhua No. 1" (red-fleshed, named B336), both of which were represented by 11-year-old grafted trees.

Methods

The Yield and Fruit Quality

The yield of the 10-year-old tree of H30-6 was measured from 2019 to 2021, and the yield of the 3-year-old grafted trees was measured in 2021. In 2019, 70 naturally set H30-6 fruits were used to calculate the seed number per loquat fruit, which was repeated three times. From 2019 to 2021, fruits of the three varieties were used each year to assess fruit quality, including average fruit weight, seed number per fruit, fruit edible rate, fruit shape index (longitudinal diameter/transverse diameter of the fruit), soluble solids (measured using a sugar refractometer), total sugar (measured using Fehling's reagent titration), reducing

sugar (measured using Fehling's reagent titration), titratable acid (measured using sodium hydroxide titration), sugar-acid ratio (total sugar/titratable acid), and vitamin C (measured using 2,6-dichlorophenol sodium indophenol). The fruit edible rate, fruit shape index, and sugar-acid ratio were calculated using the following formulas:

$$\text{Fruit edible rate } (\%) = \frac{\text{Flesh weight}}{\text{Fruit weight}} \times 100\%$$

$$\text{Fruit shape index} = \frac{\text{Longitudinal diameter per fruit}}{\text{Transverse diameter per fruit}}$$

$$\text{Sugar acid ratio } (\%) = \frac{\text{Total sugar}}{\text{Titratable acid}} \times 100\%$$

Flower Organ Observations and Quantification

The flower buds were dissected 1 day before opening (DBO) to count the flower organs and measure the stigma length. The evenness of stigma length was calculated as the ratio of the number of flowers with stigmas of the same length to the number of all flowers observed. There were 30 flowers of H30-6, H411, and B336 which were observed separately. Many of the anthers of the flowers dissected with 1 DBO were collected, dried naturally at 25°C, and stored at 4°C for subsequent experiments. We collected images of blooming flowers to show the anther dehiscence state of the three varieties in the field. Additionally, the dried anthers were observed and imaged under a Carl Zeiss Jena Axio Zoom version 16 microscope (Jena, Germany).

Pollen Morphology, Pollen Viability, and Pollen Quantity

To study pollen morphology, dried pollen grains were adhered onto the sample stage with conductive adhesive, coated with gold in a KYKY SBC-12 vacuum evaporation instrument, and observed under a Phenom ProX scanning electron microscope (Thermo Fisher Scientific, Waltham, MA, USA). For H30-6, H411, and B336, three pollen group photos were taken separately by scanning electron microscope to count the pollen with normal or abnormal shapes. The total numbers of pollen counted were 255, 159, and 297 for H30-6, H411, and B336, respectively.

There were 10 dried anthers which were placed into a 1.5 ml centrifuge tube and then crushed in 200 μl BK medium with tweezers to count the pollen in a blood counting chamber. The test was processed three times for H30-6, H411, and B336 separately.

We slightly modified Hu Shiyi's pollen viability method (Hu, 1993) to test pollen viability. The dry anthers were placed into a 1.5 ml centrifuge tube with 200 μl BK medium, and the tube was shaken to allow pollen dispersion until the pollen density was the same among H30-6, H411, and B336. Then, the anthers were removed with tweezers. Later at 45 min, 50 μl pollen suspension was transferred into a new centrifuge tube and then placed into the dark constant temperature incubator upside down at 25°C for 4 h. After culture and thorough shaking, 15 μl suspension was dropped onto a glass slide, covered with a coverslip, and observed and imaged using an Olympus (BX35) fluorescence microscope (Tokyo, Japan). At least 6 times were repeated in each material. Each repeat took at least three pictures of different visions. The germinated pollen and ungerminated pollen were counted in each repeat. The total numbers of the pollen count in the viability test were 2,743, 1,724, and 2,568 for H30-6, H411, and B336, respectively.

Anthers Development

Buds with anthers with microspore mother cell before meiosis or at meiosis stage and buds 1 DBO were collected, fixed in a 38% formalin-acetic acid-70% alcohol solution (FAA, volume ratio = 1:1:18) for 24 h, and then preserved in 70% alcohol solution at 4°C. Anthers were dissected from fixed buds for serial paraffin sections following the protocol described by Li (1987), including safranine staining, dehydration, transparency, wax immersion, and embedding. Thin serial sections of 8–10 μm were cut with a microtome (RM2235, Leica, Wetzlar, Germany). This was followed by safranine staining, dehydration, Fast Green counterstaining, and rehydration prior to examination under a light microscope (BX53, Olympus, Tokyo, Japan).

PMCs Meiosis

Anthers were dissected from buds, fixed in methyl alcohol: acetic acid solution (3:1) for 24 h, rinsed with deionized water, and then treated with 3% (w/v) cellulose, 1% (w/v) macerozyme pectinase, and 1% (w/v) snailase for 4 h. Slides were prepared using the chromosomal smear technique (Liang and Li, 1991) with Na_2HPO_4 solution prestaining for 30–40 min and 5% Giemsa staining for 10–15 min and then observed under a light microscope (BX53, Olympus, Tokyo, Japan).

Stigmas Receptivity

The benzidine staining method (Dafni et al., 2005) was used to assay stigma receptivity. The staining solution contained 1% benzidine, 3% H_2O_2, and H_2O (volumetric ratio = 4:11:22), and 3% H_2O_2 was prepared when needed. Five flowers at 1 DBO of H30-6, H411, and B336 were chosen separately, and the stigmas were revealed by removing the sepals, petals, and anthers. The stained stigmas were observed and imaged under a Carl Zeiss Jena Axio Zoom version 16 microscope (Jena, Germany).

In Situ Pollen Germination

The aniline blue staining protocol was used to test *in vivo* pollen tube growth according to Hu (1982), and the tissues were observed under a light microscope (BX53, Olympus, Tokyo, Japan). Five pollination methods were used on H30-6 to test for *in situ* pollen germination. First, natural pollination with bags was applied: some branches with many flowers were bagged before the flowers opened, and then each flower was marked on the day it started opening. More than 20 flowers were collected and fixed 2, 3, and 4 days after opening (DAO). Second, open pollination was allowed to proceed: when the flowers naturally opened, they were marked. More than 20 flowers were collected and fixed 2, 3, and 4 DAO. Third, artificial selfing was performed: some branches with many flowers were bagged before the flowers opened, and the flowers 1 DBO were emasculated and pollinated with dry pollen of H30-6. More than 20 flowers were collected 2 days after pollination (DAP) and fixed. Fourth, cross-pollination with H411 was applied: some branches with many flowers were bagged before the flowers opened, and 1 DBO, the

flowers were emasculated and pollinated with dry pollen of H411. More than 20 flowers were collected 2 DAP and fixed. Fifth, cross-pollination with B336 was conducted: several branches with many flowers were bagged before the flowers opened, and flowers at 1 DBO were emasculated and pollinated with dry pollen of B336. More than 20 flowers were collected 2 DAP and subsequently fixed.

Embryo Sac Observations

The ovules were separated from flowers 1 DBO, fixed in FAA for 48 h, and then transferred to 70% ethanol for 24 h. We followed the protocol of Zilli et al. (2015) for dehydration and removal, with the following modifications: 10% H_2O_2 (30 min), 10% H_2O_2 (2 h), 70% ethanol 2 h, 85% ethanol (2 h), 95% ethanol (2 h), 100% ethanol (2 h), 100% ethanol (overnight), methyl salicylate/ethanol 50:50 v/v (3 h), methyl salicylate/ethanol 75:25 v/v (3 h), methyl salicylate/ethanol 85:15 v/v (3 h), methyl salicylate (3 h), and methyl salicylate until observations performed through a light microscope (BX53, Olympus, Tokyo, Japan). The numbers of ovules observed in this test were 142, 289, and 123 for H30-6, H411, and B336, respectively.

Pollination Treatments

From 2017 to 2021, reciprocal crosses were performed among H30-6, H411, and B336, and pollen was collected and saved in the same way as germination in the *in situ* experiment. In the meantime, the flowers of H30-6, H411, and B336 were processed by natural pollination with bags, open pollination, and artificial selfing separately. Each year, the fruits in each treatment were harvested, and the numbers of fruits and seeds were counted. The fruit setting rate, average seed number per fruit, and seed ovule ratio were calculated using the following formulas:

$$\text{Fruit setting rate (\%)} = \frac{\text{Total fruit number of one treatment}}{\text{Total flower number of the treatment}} \times 100\%$$

$$\text{Seed ovule ratio (\%)} = \frac{\text{Total seed number of one treatment}}{\text{Flower number of the treatment} \times 10} \times 100\%$$

$$\text{Seed number per fruit (\%)} = \frac{\text{Total seed number of one treatment}}{\text{Total fruit number of the treatment}} \times 100\%$$

Chromosome Karyotype Analysis

Root tips and shoot tips were placed into 0.002 mol/L 8-hydroxyquinoline aqueous solution for 4 to 5 h in the dark, fixed in methyl alcohol: acetic acid solution (3:1, v/v) for 24 h, rinsed with deionized water, and then treated with 3% (w/v) cellulose and 1% (w/v) macerozyme for 3 h. Slides were prepared and observed as described for the meiosis slides. Chromosome karyotype analysis was according to Li and Chen (1985) standard. Chromosome types and relative lengths were determined according to Levan et al. (1964). While karyotype classification was performed according to the Stebbins (1971) standard, the asymmetrical karyotype coefficient was calculated as described by Arano (1963).

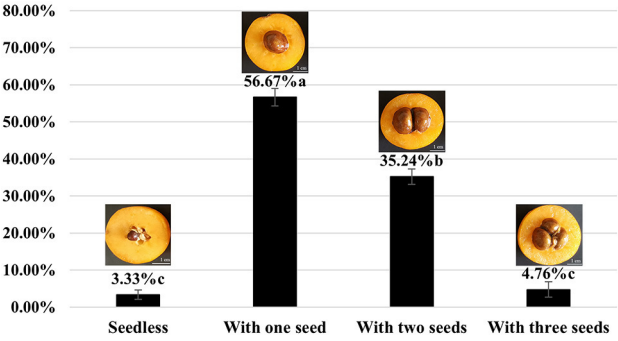


FIGURE 1 | The proportion of H30-6 open-pollinated fruits with different seed numbers. The statistical differences in means of the ratio of fruits with different seed numbers were determined using the LSD test; different letters beside the bar indicate that the means are significantly different ($P < 0.05$); values are presented as mean (± SE).

Data and Image Processing

Microsoft Excel (Redmond, Washington, USA) and SPSS 17 software (IBM, Armonk, NY, USA) was applied for statistical analysis. Adobe Photoshop CS5 (Sam Jose, USA) was applied for image processing. The parameters in the fruit quality test, the pollen quantity, normal shape pollen ratio, germination pollen ratio, fruit setting rate, seed ovule ratio, and No. of seeds per fruit were presented by mean ± standard error (SE). ANOVA for mean comparisons (LSD/Dunnett T3, $P < 0.05$) was performed. The percentage data were transformed using the arcsine square root transformation to normalization separately before the significant difference tests.

RESULTS

The Yield of H30-6 and Fruit Quality of Naturally Seated Fruit Naturally Seated Fruit

The average yield of the 10-year-old tree of H30-6 was 14.43 kg (15.19, 18.36, and 9.75 kg, from 2019 to 2021 separately), and the average yield of the 3-year-old grafted tree was 7.3 kg. The results showed that the seed numbers of open-pollinated H30-6 fruits were between 0 and 3 (**Figure 1**), and the percentage of fruits with one seed was 56.67%, followed by two-seeded fruit (35.24%). Subsequently, similar observations were made in the last 2 years, when the average seed number per fruit was 1.54, which was significantly lower than that of the controls (H411, 4.63 ± 0.22; B336, 3.23 ± 0.12). All the results proved that loquat line H30-6 was a few-seed resource. Furthermore, the successive three-year fruit quality data revealed that H30-6 had the highest fruit edible rate (70.77 ± 0.61%), and the other detection indexes were equivalent to those of the two controls (**Table 1**). Therefore, loquat H30-6 can be used as a commercial variety.

TABLE 1 | Fruit qualities during three consecutive years (2019 to 2021).

Lines/varieties	H30-6	H411	B336
Average fruit weight (g)	32.59 ± 1.47^{b}	31.58 ± 0.86^{b}	36.89 ± 0.77^{a}
Seed number per fruit	1.54 ± 0.09^{c}	4.63 ± 0.22^{a}	3.23 ± 0.12^{b}
Edible rate (%)	70.77 ± 0.61^{a}	65.3 ± 1.03^{b}	65.76 ± 0.69^{b}
Fruit shape index	1.30 ± 0.02^{a}	0.90 ± 0.01^{c}	1.11 ± 0.02^{b}
Soluble solids	9.94 ± 0.35^{b}	12.43 ± 0.56^{a}	9.28 ± 0.36^{b}
Total sugar (g/100 ml)	$7.61 \pm 0.0.13^{b}$	10.72 ± 0.59^{a}	7.66 ± 0.50^{b}
Reducing sugar (g/100 ml)	6.75 ± 0.20^{a}	6.93 ± 0.16^{a}	5.80 ± 0.17^{b}
Titratable acid (g/100 ml)	0.24 ± 0.01^{b}	0.37 ± 0.03^{a}	0.28 ± 0.01^{a}
Sugar acid ratio	31.41 ± 0.60^{a}	30.69 ± 1.29^{a}	27.39 ± 1.16^{b}
Vc (mg/100 ml)	4.65 ± 0.43^{a}	3.82 ± 0.26^{a}	3.98 ± 0.45^{a}

ANOVA for mean comparisons (LSD, $P < 0.05$) was performed among H30-6, H411, and B336 (the average fruit weight, fruit shape index, soluble solids, total sugar, and titratable acid were analyzed by Dunnett T3 test among different treatments). Lowercase letters indicate differences at the level of significance of 5%; values are presented as mean ± SE.

TABLE 2 | Flower organ numbers and evenness degree of stigma length of H30-6, H411, and B336.

Lines/ varieties	Average petal number	Average sepal number	Average anther number	Average ovule numbers	Average style number	the evenness degree of stigma length (%)
H30-6	5	5	20	10	5	80
H411	5	5	21.87	10	5	40
B336	5	5	20	10	5	80

Flower Organs

Among the flower organs, we observed the numbers of petals, sepals, anthers, and ovules, and the stigma length evenness (5, 5, 20, 10, 80%) was equal to that of B336, but H411 had more anthers (21.87) and lower stigma length evenness (40%) (**Table 2, Figures 2A–F**). Therefore, the number of flower organs is not the reason for the low seed formation of H30-6. Field observations of anther dehiscence showed that H411 and B336 released large amounts of pollen while almost no pollen was released from H30-6 anthers (**Figures 2G–I**). Similarly, the dried anthers released little pollen. These results revealed that there are abnormalities in the anthers or pollen of H30-6 (**Figures 2J–L**).

Pollen Morphology, Viability, and Quantity

The scanning electron microscopy results (**Figures 3A–F**) showed that the loquat pollen grains were medium in size, prolate, and tricolporate. We also found that the pollen sexine ornamentation of the three varieties was striated, with sparse holes between the ridges. However, there were significant differences in normal pollen ratios among H30-6, H411, and B336. The H30-6 line presented the lowest normal pollen rate (34.98%), while those of H411 and B336 were 76.91 and 92.29%, respectively (**Figures 3D–F,J**). The viability test presented similar results, as the germination rates of the H30-6, H411, and B336 lines were 4.04, 61, and 72.83%, respectively (**Figures 3G–J**). Thus, the H30-6 line exhibited low male fertility. According to Deng (2009a) standard, the H30-6 strain was a male-sterile resource, and the controls were both fertile. We also found that the pollen quantity of H30-6 was significantly lower than that of the two controls (**Figure 3K**).

Anther Anatomical Structure

The anther paraffin sections revealed no obvious structural differences among H30-6, H411, and B336 during the developmental process (**Figure 4**). However, in male gametangia dissected from flower buds 1 DBO, the controls had plump pollen (**Figures 4F,I**), while the pollen of H30-6 was distorted and shrunken (**Figure 4C**). Therefore, we hypothesize that the reason for the lower pollen release involved the pollen mother cells, not the other tissues.

PMCs Meiosis

The PMCs of B336 behaved quite normally throughout all the meiosis stages. Additionally, the normal behavior rate of H411 PMCs ranged from 78.31 to 95.64%, whereas those of H30-6 ranged from 0 to 41.03% (**Table 3, Figure 5**, **Supplementary Figure 2**). The high-frequency abnormalities of H30-6 included incorrect chromosome pairing, chromosome lag, abnormal spindle formation, unequal division, and micronucleus formation. To analyze homologous chromosome pairing, chromosome configurations at diakinesis were observed and quantified. The results showed that the average chromosome configurations were 6.87I + 9.99II + 1.07III + 0.69IV + 0.24V (H30-6), 0.70I + 16.50II + 0.02III + 0.06IV (H411), and.01I + 16.99II (B336) (**Supplementary Table 1**). From these configurations, we found that the chromosome configuration of H30-6, H411, and B336 was mainly bivalent, however, H30-6 also showed a high frequency of univalents, and 89.55% of the H30-6 PMCs presented as univalent (**Supplementary Table 1**). Due to the randomness of the bipolar spindle fibers, univalent travel to one pole will be delayed. In the later meiosis process, chromosome lag was the main phenomenon, which accounted for 65.31, 95.24, 48, 55.13, 95, and 46.32% of the observations at metaphase I, anaphase I, telophase I, metaphase II, anaphase II, and telophase II, respectively (**Table 3**, **Figure 5**). Usually, lagging chromosomes are distributed randomly, as well as losing or forming micronuclei. Some of the immature mini-microspores grew to the mature pollen stage. This was proven by the pollen images obtained by scanning electron microscopy, which showed many small pollen grains. Therefore, we believe that the high frequency of univalent chromosomes is the reason for the chromosome lag, which may cause a decrease or an increase in the inherited material, leading to an influence on the fertility of the pollen. In addition, we found that the percentages of PMCs with doubled chromosomes were as high as 16.33 and 41.03% at metaphase I and metaphase II, respectively.

Stigma Receptivity and *in situ* Pollen Germination

We observed that loquat stigmas were heart-shaped or triangular. The stigmas of H30-6, H411 and B336 were stained dark brown (**Figures 6A–C**), meaning that all the stigmas were receptive. However, the stained areas showed some differences, especially

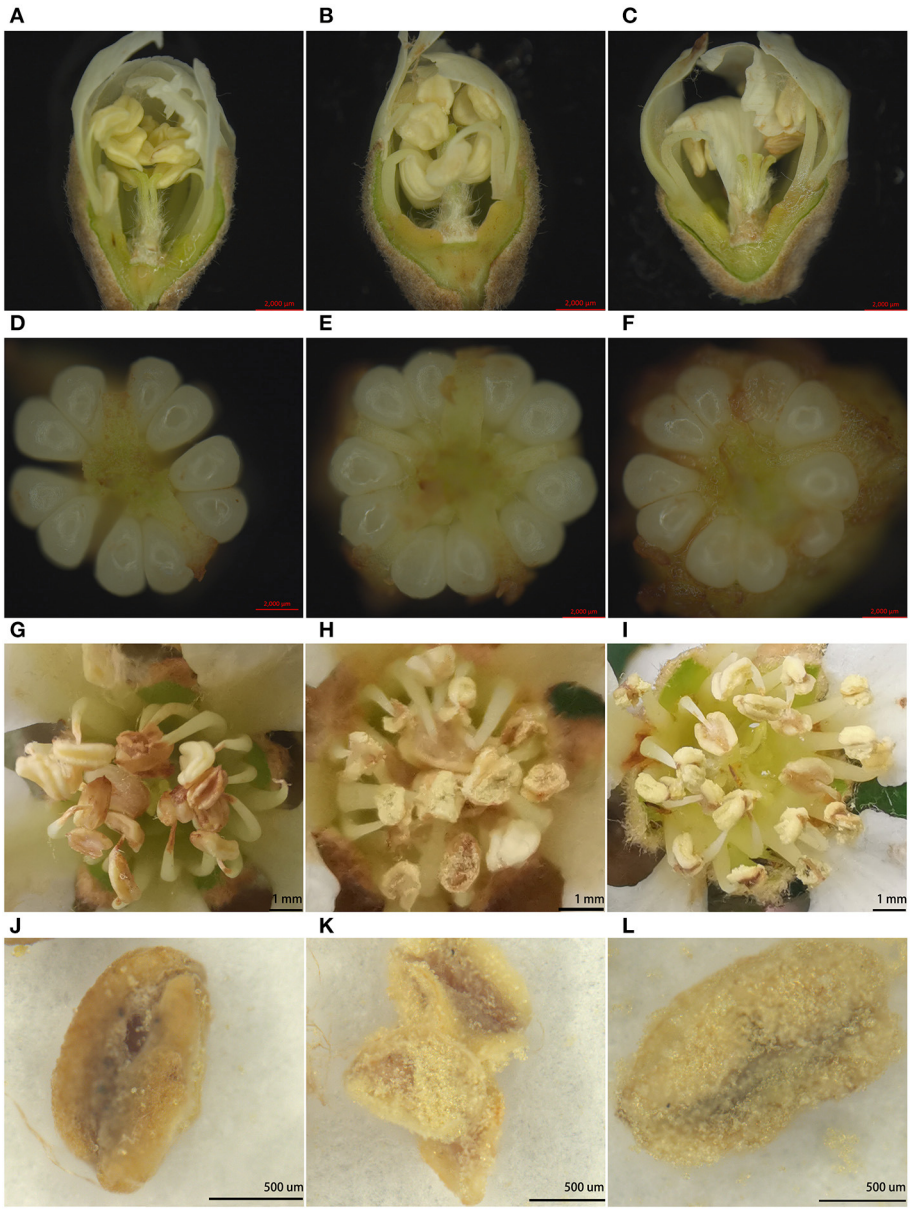


FIGURE 2 | Images of flower organs: **(A–C)**, vertical sections of flower buds 1 day before opening [H30-6 **(A)**, H411 **(B)**, B336 **(C)**]; **(D–F)**, ovules in one ovary [H30-6 **(D)**, H411 **(E)**, B336 **(F)**]; **(G–I)**, field observations of anther dehiscence [H30-6 **(G)**, H411 **(H)**, B336 **(I)**]; **(J,K)**, dried anthers imaged under a Carl Zeiss Jena Axio Zoom. v16 microscope [H30-6 **(J)**, H411 **(K)**, B336 **(L)**].

those on the stigmas of H30-6, where only the small areas of two opposite angles were stained. The aniline blue staining results showed that the pollen germinated and that the pollen tubes passed through the styles, which also indicated that the stigma receptivity was normal. We calculated the percentages of stigmas with pollen, styles with germinated pollen, and styles penetrated by pollen tubes for each of the five pollination methods. For the first method (natural pollination with bag), there was no pollen on the stigmas and no pollen germination on the styles, regardless of whether the observation was made 2, 3, or 4 DAO (**Figures 6D–F**, **Supplementary Table 2**). For the second method (open pollination), from 2 DAO to 4 DAO, the percentages of stigmas with pollen, styles with germinated pollen, and styles with pollen tubes that passed through increased (21.7, 41.18, and 46.67%; 21.71, 39.5, and 45.56%; 0, 4.2, and 16.67%; respectively) (**Figures 6G–I**, **Supplementary Table 2**). The staining results of styles pollinated with H30-6 pollen (2 DAP) showed that the percentages of stigmas with pollen, styles with germinated pollen, and styles with pollen tubes that passed through were much lower than the percentages observed under pollination with H411 and B336 pollen (**Figures 6J–L**, **Supplementary Table 2**). These results suggest that H30-6 pollen can germinate on H30-6 styles but that its low pollen viability may cause a lower *in situ* germination rate and a lower percentage of style pollen tube

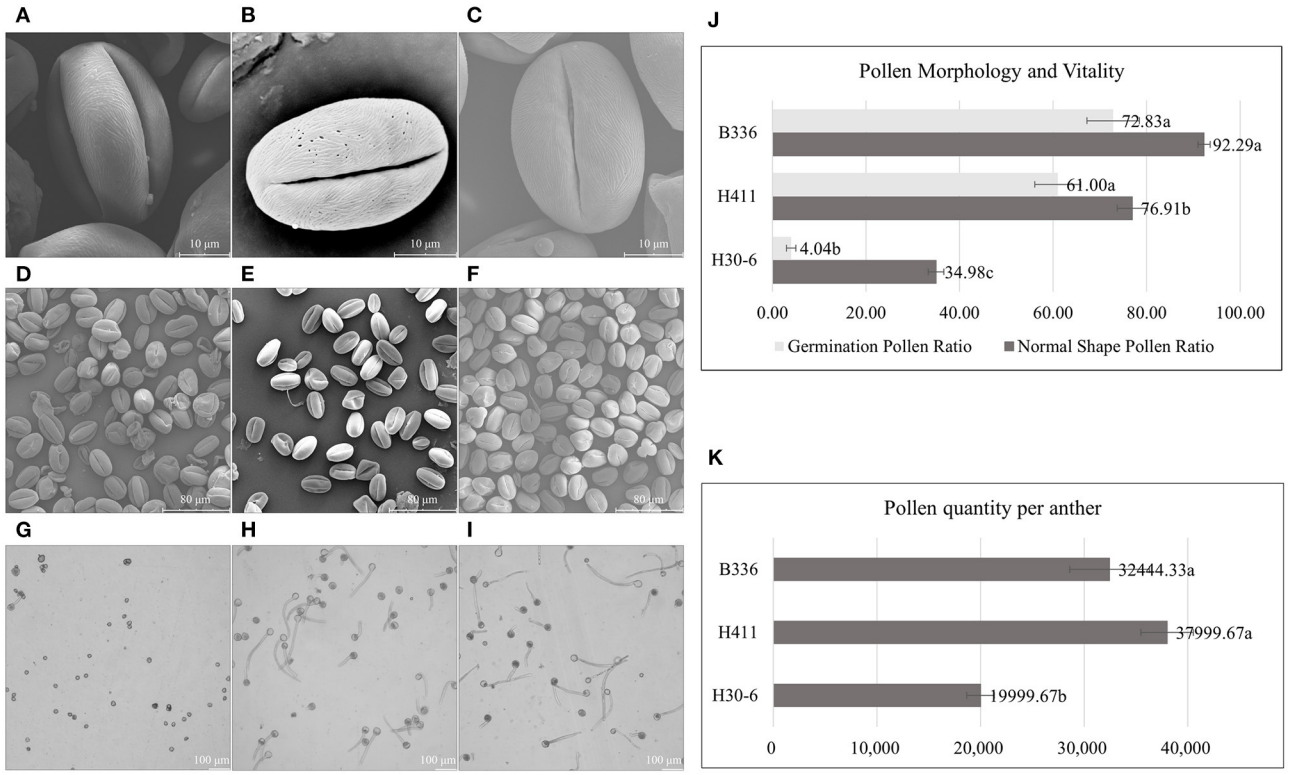


FIGURE 3 | Pollen morphology as observed *via* scanning electron microscopy, images of pollen germination, numbers of normally shaped pollen grains and germinated pollen grains, and pollen quantity per anther: **(A–C)**, equatorial views of pollen grains [H30-6 **(A)**, H411 **(B)**, B336 **(C)**] showing the sexine with striate ornamentation and sparse holes between the ridges. **(D–F)**, pollen grains of H30-6, H411, and B336 under the scanning electron microscope. **(G–I)**, *in vitro* pollen germination of H30-6, H411, and B336, respectively. **(J)** normal shape pollen ratio and germination pollen ratio of H30-6, H411, and B336. **(K)** pollen quantity per anther. The statistical differences in means of normal shape pollen ratio, germination pollen ratio among H30-6, H411, and B336 were determined using the LSD test separately; different letters beside the bar indicate that the means are significantly different ($P < 0.05$); values are presented as mean (± SE).

passage, also proving that the natural fruits of H30-6 are likely hybrid fruits.

Embryo Sac Structure

As we observed ovules at 1 DBO, the embryo sacs should have been mature with 8 nuclei. H30-6, H411, and B336 all had normal embryo sacs, but the normal embryo sac rate of H30-6 was only 22.54%, which was obviously lower than that of the controls (H411 71.79%, B336 95.93%) (**Figure 7**, **Supplementary Table 3**). Each ovary had 10 ovules, so in H30-6, the normal ovule number per ovary was 2.25. This number was close to the average seed number of H30-6 natural fruits (1.54), so we speculate that embryo sac abnormality may be the most restrictive factor for lower seed formation. The abnormal embryo sacs could be divided into three main kinds: those remaining in the tetrad stage (with one functional macrospore or with four macrospores, 17.27%), those remaining in the binucleate embryo sac stage (1.82%), and those without an embryo sac (68.18%) (**Figure 7**, **Supplementary Table 3**). We hypothesize that majority of functional macrospores were disabled, which caused the differentiation shutoff or degradation of female gametes.

Fruit and Seeds Obtained Through Various Pollination Treatments

The pollination treatments performed over consecutive years showed that when artificial selfing was applied to H30-6, the average fruit setting rate, average seed number per fruit and average seed ovule ratio were 9.93 ± 6.08%, 1.82 ± 0.28%, and 2.11 ± 1.44%, respectively (**Table 4**). Therefore, we believe that H30-6 is a self-compatible line. H30-6 pollination with H411 or B336 pollen dramatically improved the fruit setting rate (25.2 ± 6.23% and 26.68 ± 6.37%, respectively) compared with those under artificial selfing (9.93 ± 11.64% and 0.47 ± 2.94%) and natural pollination with bagging (2.13 ± 5.46% and 0.09 ± 1.24%) (**Table 4**). When H30-6 was open-pollinated, the fruit setting rate was significantly increased than those under natural pollination with bagging. These results showed the low fertility of H30-6 pollen was a limiting factor for the fruit setting. However, all pollination treatments to H30-6 showed no significant difference in seed number per fruit (from 1.59 ± 0.14 to 2 ± 0.17), which proved that the male gametes were not the main reason for the low seeds' formation in the natural fruits of H30-6 and may indicate that at least 1 to 2 seeds were usually needed for the loquat fruit set of H30-6

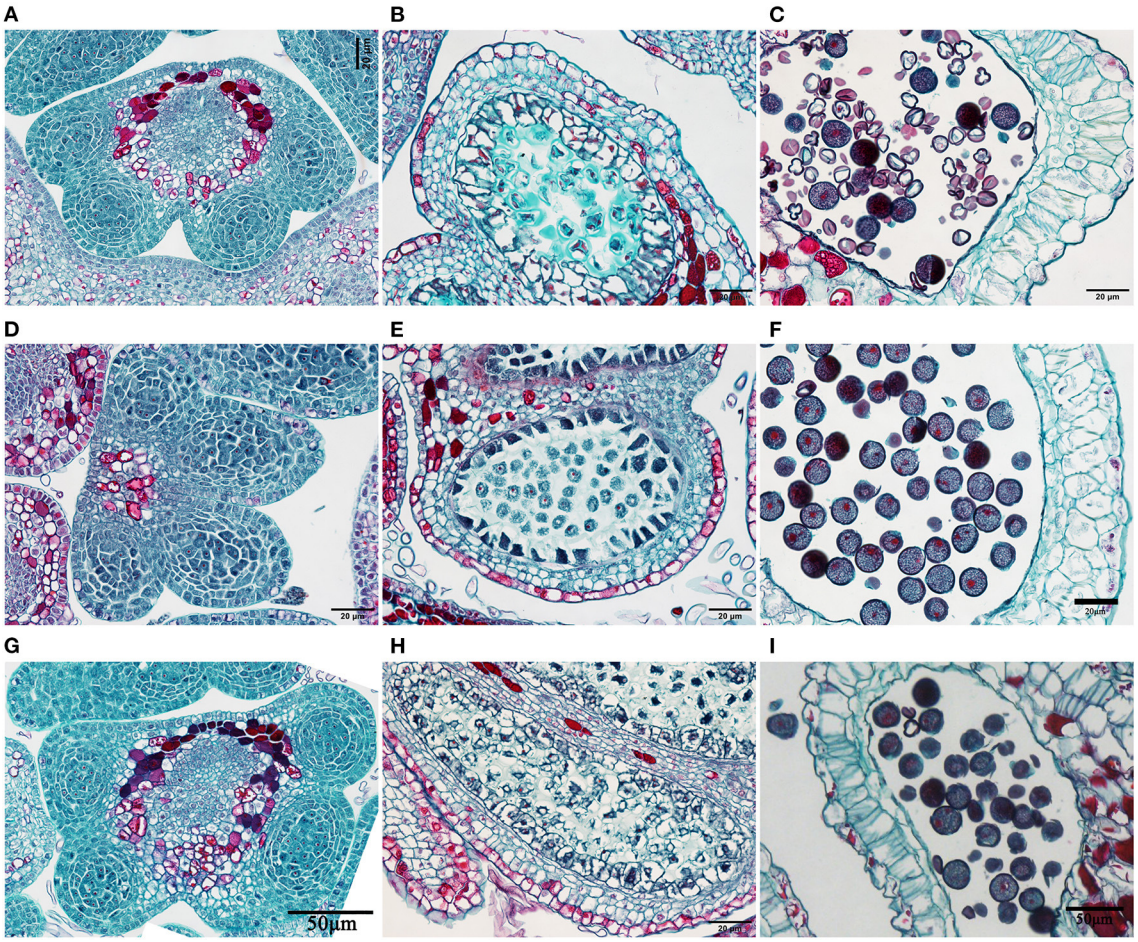


FIGURE 4 | Images of anther slices: **(A–C)**, Anther of H30-6 at the microspore mother cell stage (with microspore mother cells before meiosis), tetrad stage of meiosis stage, as well as the mature pollen stage; **(D–F)** anther of H411 at the microspore mother cell stage, meiosis stage, and mature pollen stage; **(G–I)** anther of B336 at the microspore mother cell stage, tetrad stage, and mature pollen stage.

(**Table 4**). The average seed number per fruit was also close to the former results of the number of ovules with normal embryo sac (2.25) in the ovary of H30-6. All these results proved that embryo sac abnormality was the most important restrictive factor for seed formation of H30-6. The seed ovule ratios were dramatically increased when H30-6 was pollinated with H411 or B336 pollen than that natural pollinated with the bag. But there was no significant difference in seed ovule ratio among the four pollination treatments, pollination with pollen of H411, pollination with pollen of B336, open pollination, and artificial selfing when H30-6 was the female parent (**Table 4**). Meanwhile, the fruit setting rate and seed ovule ratio were without significant difference among pollination treatments (pollinated with pollen of B336 or H30-6, and artificial selfing) to H411 (**Table 4**). There was no significant difference in fruit setting rate and seed ovule ratio among pollination treatments (pollinated with pollen of H30-6 or H411, and artificial selfing) to B336 (**Table 4**). The results above showed that H30-6 can be used as a male parent in loquat breeding.

Somatic Cell Karyotypes

The small amount and low vitality of pollen and the differentiation shutoff or degradation of the female gametes showed that H30-6 had low gametophyte fertility, which caused the low seed formation of the loquat fruit. We believe that there must be some common reasons for the low gametophyte fertility. As the abnormal meiosis of PMCs caused the male gametes to be sterile and the disabled functional macrogametes came from the MMCs, we observed the chromosomes of H30-6 to try to determine a common underlying factor.

The chromosome number shoot and stem-tip cells showed that H30-6 was diploid ($2n = 2x = 34$), which was the same as that of the controls; thus, the univalent was observed during PMC meiosis was not caused by an unbalanced chromosome number. Then, we analyzed the chromosome karyotype. The relative length (%) (long arm + short arm = full length), arm ratio (long/short) and chromosome type were assessed (**Supplementary Table 4**). The karyotype formulas were as follows: $2n = 2x = 26m + 8sm$ (H30-6), $2n =$

TABLE 3 | Counting of pollen mother cell (PMC) meiosis phenomena.

Meiosis stage	Phenomenon	Lines/varieties					
		H30-6		H411		B336	
		Number of PMCs	Percentage (%)	Number of PMCs	Percentage (%)	Number of PMCs	Percentage (%)
Metaphase I	Normal	10	20.41	490	81.53	135	97.82
	With chromosomes lag	32	65.31	104	17.3	3	2.17
	Chromosomes doubled	8	16.33	7	1.16		
	Total number of PMCs	49		601		138	
Anaphase I	Normal	0	0	277	90.52	112	100
	With chromosomes lag	20	95.24	24	7.84		
	With micronucleus	8	38.1				
	With chromosomes bridge			5	1.63		
	Total number of PMCs	21		306		112	
Telophase I	Normal	7	9.33	329	95.64	58	100
	With chromosomes lag	36	0.48	15	4.36		
	With chromosomes bridge	1	1.33				
	With micronucleus	32	42.67				
	Total number of PMCs	75		344		58	
Prophase II	Normal	15	31.91	255	93.41	219	100
	With dissociative chromosomes	17	36.17				
	With micronucleus	12	25.53	18	6.59		
	Unequal separation of chromosomes	2	4.26				
	Chromosomes doubled	1	2.13				
	Total number of PMCs	47		273		219	
Metaphase II	Normal	5	3.21	261	84.74	148	98.01
	With chromosomes lag	86	55.13	43	13.96	3	1.99
	Chromosomes doubled	64	41.03				
	With micronucleus	46	29.49	4	1.3		
	With perpendicular spindle	6	3.85				
	With splayed spindle	9	5.77				
	Total number of PMCs	156		308		151	
Anaphase II	Normal	0	0	195	78.31	86	98.85
	With chromosomes lag	57	91.94	15	6.02	1	1.15
	With micronucleus	21	33.87	9	3.61		
	With chromosomes bridge			30	12.05		
	With splayed spindle	8	12.9				
	With perpendicular spindle	2	3.23				
	With cross spindle	1	1.61				
	Total number of PMCs	62		249		87	
Telophase II	Normal	17	7.36	299	81.25	580	96.99
	With chromosomes lag	107	46.32			13	2.17
	Unequal separation of chromosomes	50	21.65	3	0.82		
	With micronucleus	97	41.99	38	10.33	2	0.33
	Dyad			2	0.54		
	Triad	11	4.76	14	3.8	3	0.5
	Polyad	5	2.16	12	3.26		
	Total number of PMCs	231		368		598	

$2x = 18\,m + 16$ sm (H411), and $2n = 2x = 30\,m + 4$ sm (B336). The chromosome lengths varied from 1.54 ± 0.08-$2.63 \pm 0.09\,\mu m$ (H30-6), 2.02 ± 0.16–$3.59 \pm 0.16\,\mu m$ (H411), and 1.57 ± 0.05–$2.5 \pm 0.1\,\mu m$ (B336). According to the chromosome field theory, the chromosomes belonged to the small chromosome. The ratios of the longest chromosome to the shortest were 1.71 (H30-6), 1.78 (H411), and 1.59 (B336), and chromosome arm ratios above 2 accounted for 0.03 (H30-6),

0.13 (H411), and 0.06 (B336) of the total ratios. Therefore, according to Stebbins's (1971) classification, the karyotypes were all 2A types. The asymmetrical karyotype coefficients were 59.03% (H30-6), 61.95% (H411), and 59.23% (B336). **Figure 8** shows the chromosome images, karyograms, and idiograms of H30-6, H411, and B336, respectively. The chromosome observation results revealed no obvious chromosomal numerical or structural abnormalities.

DISCUSSION

Low Female Gametophyte Fertility Is Mainly Responsible for the Low Seed Formation of H30-6

Few-seed usually results from low gametophyte fertility (Xiao et al., 2007; Rédei, 2008; Qiu et al., 2012; Li et al., 2016; Royo et al., 2016), self-incompatibility (Zhang et al., 2012; Gambetta et al., 2013; Garmendia et al., 2019) and embryo abortion (Pearson, 1932; Tang, 2001). There was no self-incompatibility found for H30-6, which can be demonstrated proofed by the result of artificial selfing and *in situ* pollen germination. As a result, low fertility should be responsible for the few-seed nature of H30-6. There were no abnormalities found in flower morphology, but only limited pollen existed on the dehiscent anther indicating a low male gamete fertility character. Additionally, this was verified according to the vertical sections of mature anther and pollen morphology observation. Further, the pollen viability test showed that only 4.04% of pollen could germinate *in vitro*. So H30-6 was a male sterile material depending on Deng (2009a) standard. However, the pollination treatment results showed that there was no significant difference in the seed number (1.59-2) per fruit when the pollen of H30-6, H411, and B336 was used or open-pollinated. Meanwhile, the former embryo sac observation results showed that there were only 2.25 normal ovules in each ovary of H30-6. It was reasonable to assume that abnormal embryo sacs should be responsible for low seed formation.

May Abnormal Meiosis Lead to Low Female Gametophyte Fertility of H30-6?

The formation of both male and female gametophytes is a precisely coordinated process, including the formation of gametophyte mother cells, meiosis, and gametophyte development. Therefore, there are three likely mechanisms leading to low gametophyte fertility: pre meiosis abnormality, such as a lack of sporogenous cells (Schiefthaler et al., 1999; Yang et al., 1999) and defects in meiotic entry (Hong et al., 2012), meiosis abnormalities, such as arrests at one meiosis period (Siddiqi et al., 2000) and aberrant chromosome distribution (Hoyt and Geiser, 1996; d'Erfurth et al., 2008; De Storme and Geelen, 2011), and postmeiotic abnormality, especially reproductive companion cell abnormalities, which will cause gametophytes to be sterile (Chaubal et al., 2000; Guo et al., 2001; Suzuki et al., 2001; Smith et al., 2002; Kawanabe et al., 2006; Li et al., 2006; Zhang et al., 2006). In this study, we found no premeiotic or post-meiotic abnormalities but a multitude of meiosis abnormal phenomena of PMCs. Therefore, low male fertility is caused by abnormal meiosis. Although we did not observe the meiosis of megasporocytes, the differentiation shutoff or degradation of female gametes showed that most of the functional macrospores were disabled, which we speculated could be due to that there were meiotic abnormalities in megasporocytes.

Possible Mechanism of H30-6 Abnormal Meiosis

Chromosome number and structural changes can lead to abnormal meiosis (Grandont et al., 2014; Rocha et al., 2019). Low fertility is often found in polyploids, which have been used for loquat seedless and few-seed breeding (Maria et al., 2013). H30-6 was identified as being a diploid line. That was to say, the abnormal meiosis of H30-6 was not caused by polyploidization. A recent report showed that a 2.09 Mb chromosome fragment translocation causes abnormalities during meiosis and leads to pollen semisterility and fewer seeds in watermelon (Tian et al., 2021). Additionally, the fragment translocation results in one quadrivalent ring of four chromosomes at prometaphase I during meiosis. During the meiosis of H30-6, the earliest abnormalities were found during diakinesis, and a high frequency of univalent chromosomes (6.87) was the main abnormality. The most likely reason was that homolog pairing failure caused abnormal meiosis.

Homolog pairing is subdivided into three levels: preselection of homologous chromosomes, homologous connections by the synaptonemal complex (the process also known as synapsis), and the precise matching of DNA molecules in the process of genetic exchange and conversion (the process also known as gene recombination) (Moens, 1969; Loidl, 1990). Abnormalities occurring in all three of these levels may result in the univalent characteristics found in diakinesis.

When or how homologous chromosomes select each other is still unknown (Bozza and Pawlowski, 2008); However, chromosomal movements that are performed prior to and as a precondition to any homologous interaction have been widely supported (Loidl, 1990; Dawe et al., 1994; Bass, 1997), especially when considering telomere mediated movement (Lange, 1998; Nimmo et al., 1998; Trelles-Sticken and Scherthan, 2000; Sepsi and Schwarzacher, 2020). Varas et al. (2015) reported that two Sad1/UNC-84 (SUN)-domain proteins (AtSUN1 and AtSUN2) play an important role in telomeres movement during meiosis, and the double mutant exhibited severe meiotic defects, including a delay in the progression of meiosis, an absence of full synapsis, and a reduction in the mean cell chiasma frequency.

The synaptonemal complex stabilizes the presynaptic alignment of the axes of the homologous chromosomes and promotes the maturation of crossover recombination events, which is widely conserved in sexually reproducing eukaryotes with the lateral elements (LEs), central elements (CEs), and transverse element (transverse filaments, TFs) (Gillies and Moens, 1984; Christa, 1996; Dawe, 1998). Thus far, several proteins involved in LE, such as budding yeast Hop1 and Red1;

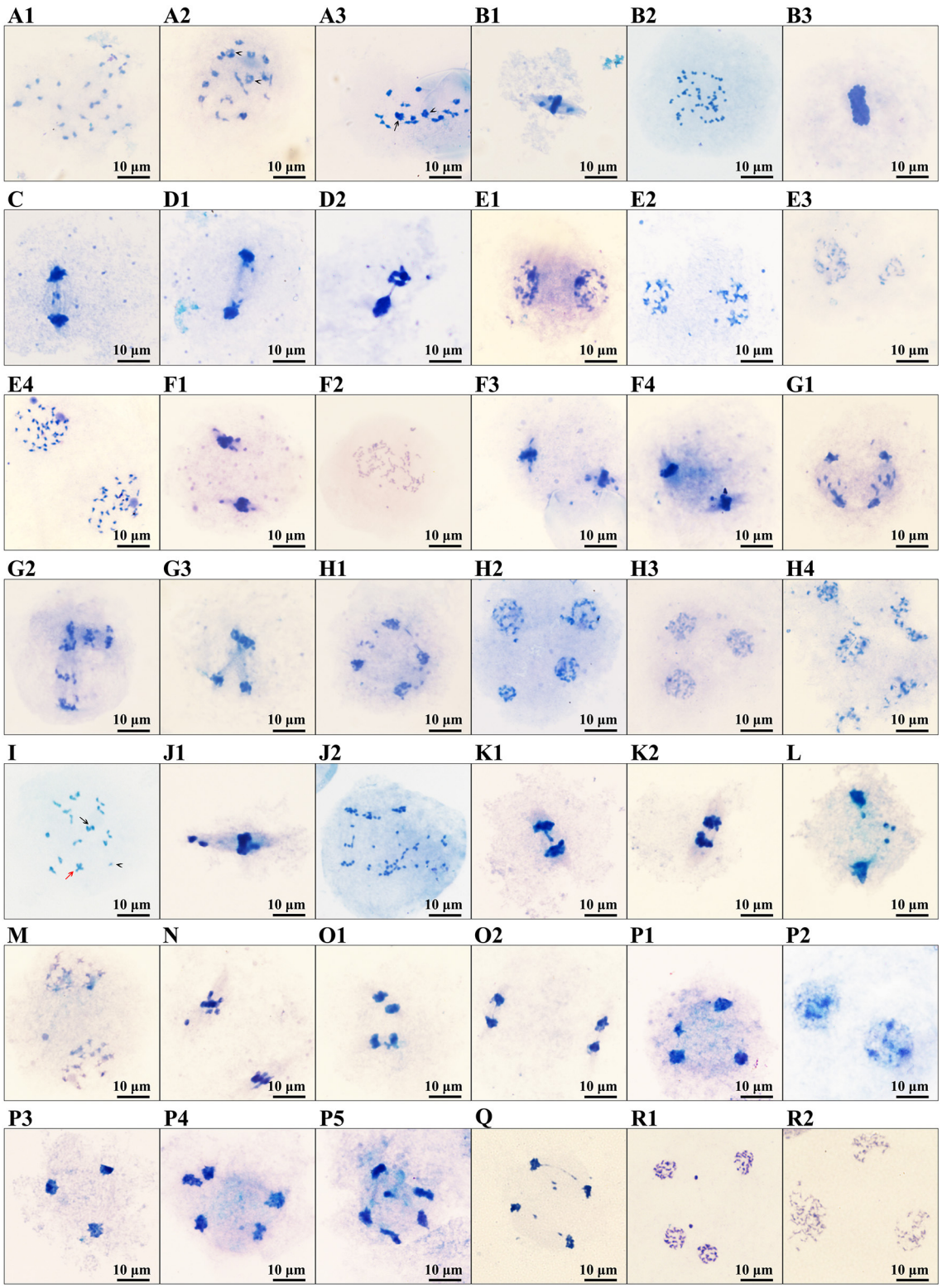


FIGURE 5 | Images of abnormal phenomena of PMCs meiosis: **(A1–H4)**, images of PMCs meiosis of H30-6: **(A1–A3)**, diakinesis, **(A1)** showing the univalent, **(A2)** showing the trivalent, **(A3)** showing the quadrivalent (arrowhead) and pentavalent (arrow); **(B1–B3)**, metaphase I, **(B1)** showing a chromosomes lag, **(B2,B3)** showing doubled chromosomes; **(C)** showing chromosomes lag and micronucleus in anaphase I; **(D1,D2)**, Telophase I, showing chromosomes lag, micronucleus

(Continued)

FIGURE 5 | and the bridge of chromosomes; **(E1–E4)**, prophase II, **(E1)** showing the dissociative chromosome, **(E2)** showing the micronucleus, **(E3)** showing the unequal separating of chromosomes, **(E4)** showing the doubled chromosomes; **(F1–F4)**, metaphase II, **(F1)** showing chromosomes lag and micronucleus, **(F2)** showing doubled chromosomes, **(F3)** showing perpendicular spindle, **(F4)** showing splayed spindle; **(G1–G3)** anaphase II, **(G1)** showing the splayed spindle chromosomes lag, **(G2)** showing perpendicular spindle and chromosomes lag, **(G3)** showing cross spindle and micronucleus; **(H1–H4)**, telophase II, **(H1)** showing chromosomes lag and micronucleus, **(H2)** showing unequal separation of chromosomes, **(H3)** showing triad and micronucleus, **(H4)** showing polyad. **(I–P5)**, images of PMCs meiosis of H411: **(I)** showing the univalent (arrowhead), trivalent (arrow), and quadrivalent (red arrow); **(J1,J2)**, metaphase I, **(J1)** showing chromosomes lag, **(J2)** showing doubled chromosomes. **(K1,K2)**, anaphase I, **(K1)** showing chromosome lag and chromosome bridge, **(K2)** showing chromosomes bridge; **(L)** showing chromosomes lag and micronucleus in telophase I; **(M)** showing micronucleus in prophase II; **(N)** showing chromosomes lag in metaphase II; **(O1,O2)**, anaphase II, **(O1)** showing chromosomes lag and chromosomes bridges, **(O2)** showing chromosome bridge; **(P1–P5)** telophase II, **(P1)** showing unequal separating of chromosomes, **(P2)** showing dyad, **(P3)** showing triad and micronucleus, **(P4,P5)** showing polyad and micronucleus. **(Q–R2)**, images of PMCs meiosis of B336: **(Q)** showing chromosome lad and chromosomes bridge in anaphase II; **(R1)** showing micronucleus, **(R2)** showing a triad. Scale: 10 μm.

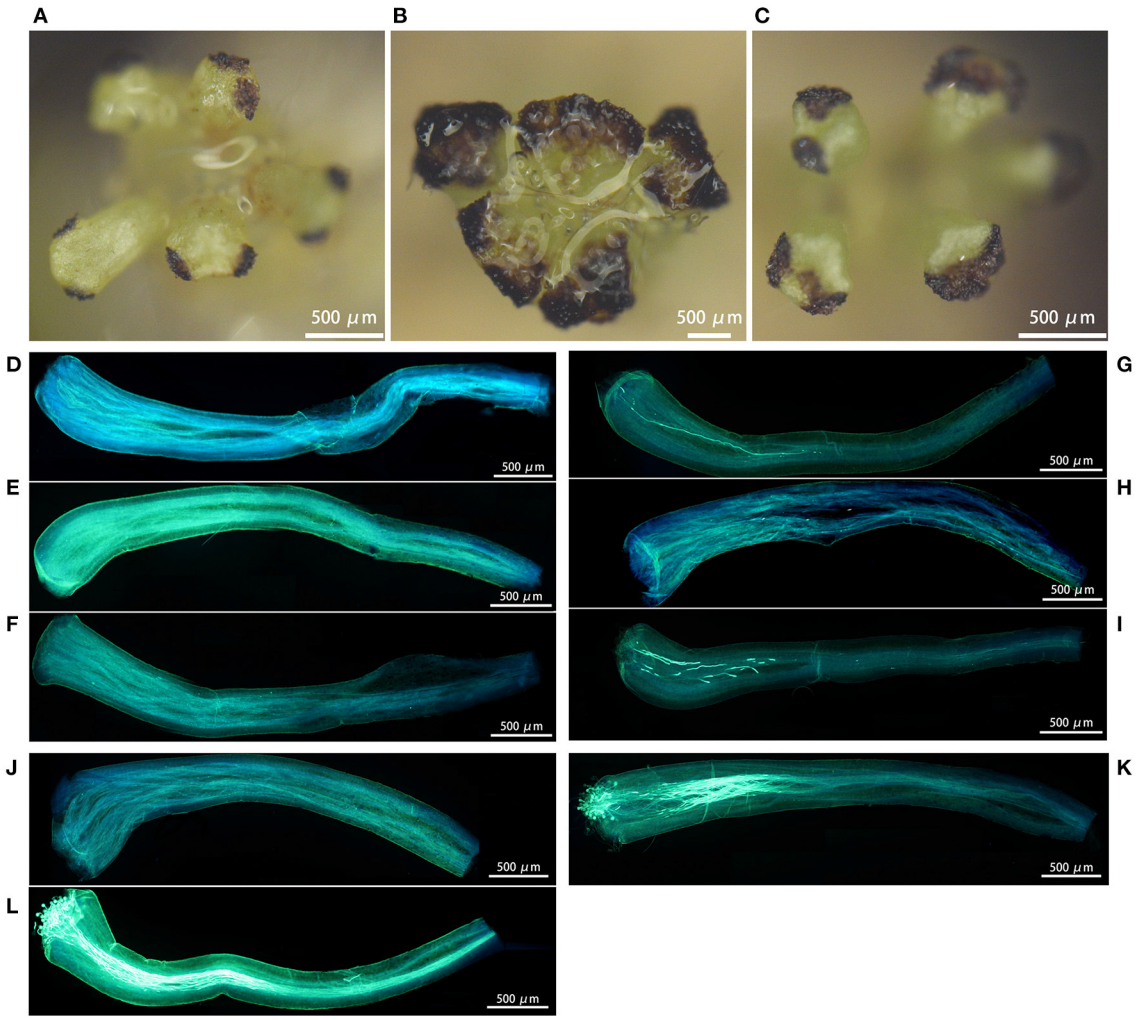


FIGURE 6 | Images of the stigmas receptivity test and pollens *in situ* germination: **(A–C)** the stained area showing the stigmas receptivity of H30-6, H411, and B336 respectively. **(D–F)** showing pollen *in situ* germination at 2, 3, and 4 days after H30-6 bagging natural pollination receptivity. **(G–I)** showing pollen *in situ* germination at 2, 3, and 4 days after H30-6 open pollination receptivity. **(J–L)** showing pollen *in situ* germination at 2 days after H30-6 artificial selfing, pollination with pollens of H411, and pollination with pollens of B336.

mice SYCP2 and SYCP3; nematode HIM-3, HTP-1, HTP-2, and HTP-3; and plants proteins ASY1/PAIR2 and ASY3/PAIR3 have been reported (Hollingsworth et al., 1990; Smith and Roeder, 1997; Yang et al., 2006; Sanchez-Moran et al., 2008; Ferdous et al., 2012; Lui and Colaiacovo, 2013). Furthermore, yeast Zip1, mice SYCP1, Drosophila C (3) G, nematode SYP-1/2/3/4, Arabidopsis thaliana ZYP1, and rice ZEP1 have been reported to be components of TFs (Sym et al., 1993; Page and Hawley, 2001; De Vries et al., 2005; Higgins et al., 2005; Wang et al., 2010; Lui and Colaiacovo, 2013).

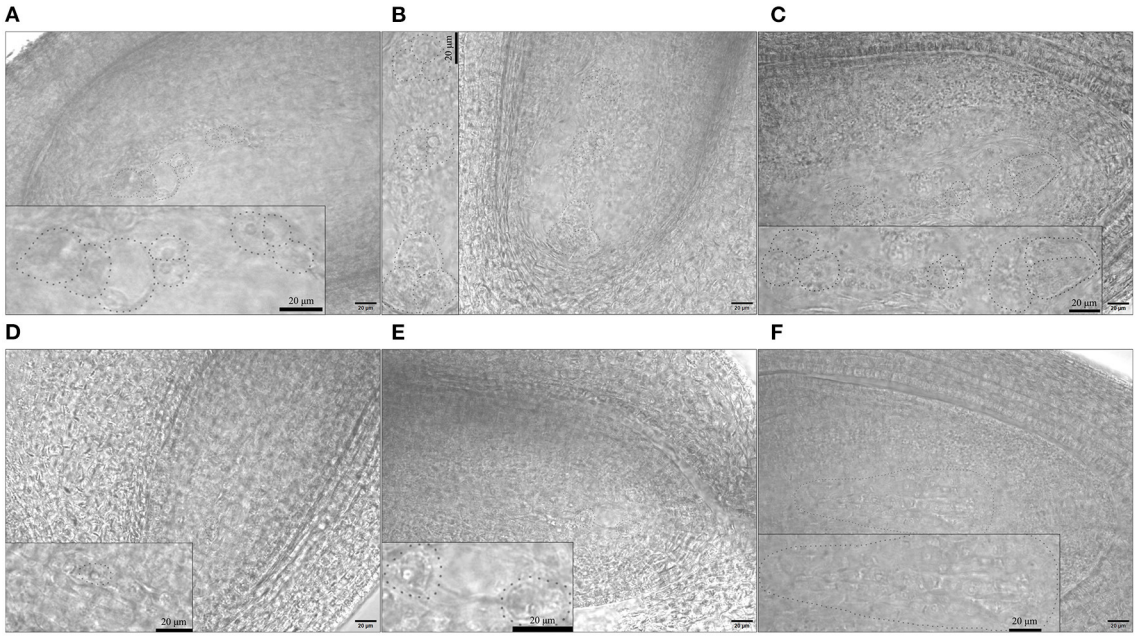


FIGURE 7 | Images of normal and abnormal embryo sacs: **(A–C)** normal embryo sacs of H30-6, H411, and B336, respectively. **(D–F)** abnormal embryo sacs of H30-6: **(D)** embryo sac at the tetrad stage, **(E)** embryo sac at the binucleate embryo sac stage, **(F)** without embryo sac. The embryo sac area was magnified in the small box on each image. Scale: 20 μm.

TABLE 4 | The continuous years of pollination treatments results (2018 to 2021).

Female parent (♀)	Pollination treatment	No. of flowers	No. of fruits	No. of seeds	Fruit setting rate (%)	No. of seeds pre fruit	Seed ovule ratio
H30-6	×H411 (♂)	903	196	360	25.20 ± 6.23^{a}	1.90 ± 0.15	4.64 ± 1.04^{a}
	×B336 (♂)	904	204	400	26.68 ± 6.37^{a}	2.00 ± 0.17	5.60 ± 1.73^{a}
	Artificial selfing	887	96	200	9.93 ± 6.08^{bc}	1.82 ± 028	2.11 ± 1.44^{ab}
	Open pollination	720	159	261	14.20 ± 3.84^{ab}	1.59 ± 0.14	2.28 ± 0.73^{ab}
	Bagging natural pollination	1,203	6	11	0.47 ± 0.35^{c}	1.88 ± 0.13	0.09 ± 0.06^{b}
H411	×H30-6 (♂)	762	144	547	19.06 ± 8.94^{ab}	3.22 ± 0.72	7.29 ± 4.70^{ab}
	×B336 (♂)	780	222	874	28.22 ± 10.63^{ab}	3.63 ± 0.40	11.08 ± 5.39^{ab}
	Artificial selfing	943	287	1,187	31.56 ± 10.83^{a}	3.89 ± 0.60	12.93 ± 5.79^{a}
	Open pollination	585	98	346	13.15 ± 4.26^{ab}	3.39 ± 0.49	4.90 ± 2.46^{ab}
	Bagging natural pollination	1219	81	224	8.40 ± 3.73^{b}	2.62 ± 0.33	2.41 ± 1.15^{b}
B336	×H30-6 (♂)	731	167	574	20.68 ± 10.53^{ab}	2.66 ± 0.70	6.96 ± 4.64^{ab}
	×H411 (♂)	706	256	932	35.70 ± 7.45^{ab}	3.62 ± 0.63	13.29 ± 4.03^{ab}
	Artificial selfing	618	260	953	53.20 ± 20.23^{a}	3.47 ± 0.41	18.71 ± 6.86^{a}
	Open pollination	615	104	302	12.33 ± 1.77^{b}	3.06 ± 0.52	3.55 ± 1.16^{b}
	Bagging natural pollination	1,216	60	138	10.86 ± 7.95^{b}	2.22 ± 0.15	2.55 ± 1.95^{b}

ANOVA for mean comparisons (LSD, $P < 0.05$) was performed among different treatments of the same female parent (the seed ovule ratio was analyzed by Dunnett T3 test among different treatments when H30-6 was a female parent). Lowercase letters indicate differences at the level of significance of 0.05; values are presented as mean ± SE.

The gene recombination structure, which is also known as the visualized crossover (CO), links the homologs after diplotene in most organisms. Typically, every chromosome undergoes at least one obligate CO event to ensure proper segregation at metaphase I (Jones, 1984). The double-strand break repair model (DSBR) explained the gene recombination (Szostak et al., 1983; Mercier et al., 2015). Meiotic DNA DSBs are formed by the highly conserved SPO11 protein (Edlinger and Schlogelhofer, 2011; Bernard, 2013) and PRD1, PRD2, AtPRD3/OsPAIR1, DFO, and CRC1 proteins in plants (De Muyt et al., 2009; Miao et al.,

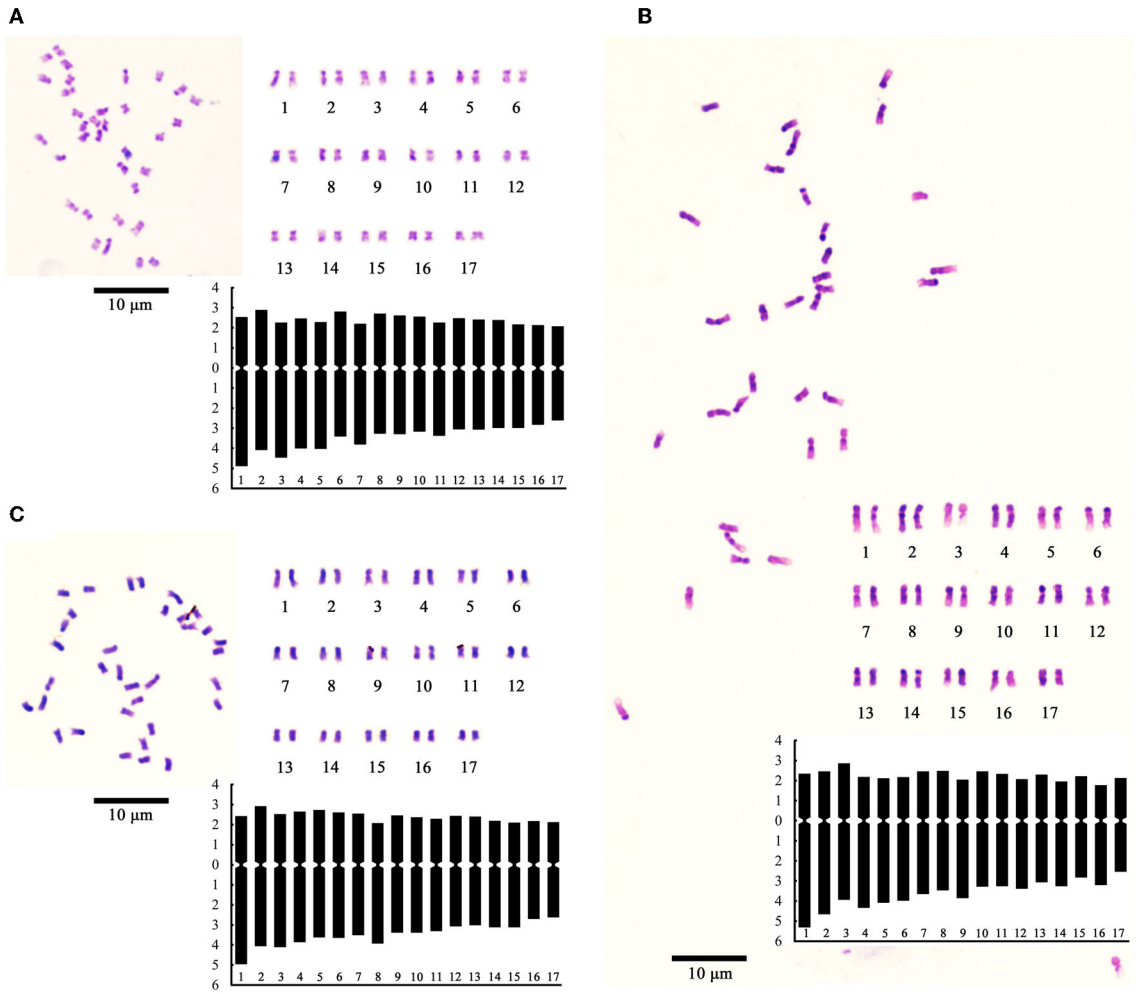


FIGURE 8 | Images of chromosome photo, karyogram, and ideogram: **(A–C)** showing chromosome images, karyograms, and idiograms of H30-6, H411, and B336 respectively.

2013). In most species, homologous chromosomes do not pair properly in mutants that do not normally form meiotic DNA double-strand-breaks (DSBs) (Peoples-Holst and Burgess, 2005; Yu et al., 2010; Ji et al., 2016; Robert et al., 2016). The plant MRE11, RAD50, and COM1 genes are functionally conserved in the early steps of DSB processing, and their respective mutants all display strong chromosome fragmentation at anaphase I (Mercier et al., 2015). In Arabidopsis, DMC1-mediated inter-homologous (IH) DNA repair is the predominant pathway of DSB repair, and when DMC1 is missing, RAD51 works as a backup pathway leading to a complete absence of synapsis and bivalent formation (Couteau et al., 1999; Deng and Wang, 2007). Finally, the CO formation relies on two pathways: the ZMM pathway, which relies on a group of proteins called ZMMs, as well as proteins MLH1 and MLH3, and the non-ZMM CO pathway, in which protein MUS81 has been characterized in the non-ZMM CO pathway in plants (Mercier et al., 2015). The mutants of the ZMM pathway and MUS81 were reported to exhibit various degrees of decrease in COs (Mercier et al., 2015).

Therefore, to clarify the molecular mechanism of H30-6 abnormal meiosis, the above proteins involved in the homolog pairing constitute the main points of this study.

CONCLUSIONS

The diploid loquat line H30-6 is a potentially commercial few-seed (1.54 per fruit) loquat, with high fruit edible rate (70.77%). The diploid loquat line H30-6 is a male sterile material, and this sterility is caused by the abnormal meiotic synapses. The low fertility of H30-6 pollen was a limiting factor for the fruit setting. The main reason for the few seeds character of H30-6 was the female gamete abnormality.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

GL, QG, and JD conceived and designed the research and revised the manuscript. QX performed the experiments and wrote and revised the manuscript. XW and SX helped with the data analysis and revised the manuscript. PW and SL managed the material collection. XL and QH helped with the material filed management. HS, DW, YX, SW, and DJ helped with the fund management, drug purchase, and instrument management. All the authors read and approved the final manuscript.

FUNDING

This research was funded by the National Key R&D Program of China (No. 2019YFD1000900), the National Nature Science Foundation of China (32171820), the Chongqing Science and Technology Commission (cstc2021jscx-gksbX0010 and cstc2021jcyj-sxmX1156), the Innovation Research Group Funds for Chongqing Universities (CXQT19005), and the Chongqing Forestry Administration (YuLinKeYan2022-14).

ACKNOWLEDGMENTS

QX expresses her great and deep appreciation to all the laboratory members for so much help and so many great moments. We also thank the reviewers and the editor for the helpful comments and suggestions on this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.882965/full#supplementary-material

REFERENCES

Arano, H. (1963). Cytological studies in subfamily Carduoideae (compositae) of Japan IX. *Bot. Mag. Tokyo*. 76, 32–39. doi: 10.15281/jplantres1887.76.32

Bass, H. W. (1997). Telomeres cluster *de novo* before the initiation of synapsis: a three-dimensional spatial analysis of telomere positions before and during meiotic prophase. *J. Cell Biol*. 137, 5–18. doi: 10.1083/jcb.137.1.5

Bernard, D. M. (2013). Initiation of meiotic recombination: how and where? conservation and specificities among eukaryotes. *Annu. Rev. Genet*. 47, 563–599. doi: 10.1146/annurev-genet-110711-155423

Bozza, C. G., and Pawlowski, W. P. (2008). The cytogenetics of homologous chromosome pairing in meiosis in plants. *Cytogenet. Genome Res*. 120, 313–319. doi: 10.1159/000121080

Chaubal, R., Zanella, C., Trimnell, M. R., Fox, T. W., and Bedinger, A. P. (2000). Two male-sterile mutants of Zea Mays (Poaceae) with an extra cell division in the anther wall. *Am. J. Bot*. 87, 1193–1201. doi: 10.2307/2656657

Chen, W., Feng, J., Qin, Q., Liu, X., Wu, J., and Xie, M. (2006). Characteristics of sugar metabolism and accumulation in GA_3 induced parthenocarpic white flesh Loquat 'Ninghai Bai' Fruit. *Acta Hortic. Sin*. 33, 471–476. doi: 10.16420/j.issn.0513-353x.2006.03.004

Christa, H. (1996). Synaptonemal complexes: structure and function. *Curr. Opin. Cell Biol*. 8, 389–396. doi: 10.1016/S0955-0674(96)80015-9

Couteau, F., Belzile, F., Horlow, C., Grandjean, O., Vezon, D., and Doutriaux, M. P.. (1999). Random chromosome segregation without meiotic arrest in both male and female meiocytes of a *dmc1* mutant of *Arabidopsis*. *Plant Cell*. 11, 1623–1634. doi: 10.1105/tpc.11.9.1623

Dafni, A., Kevan, P. G., and Husband, B. C. (2005). *Practical Pollination Biology*. Cambridge: Enviroquest Ltd. p, 120–121.

Dang, J., Guo, Q., Xiang, S., He, Q., Sun, H., Wu, D., et al. (2019a). A New Red-fleshed and Seedless Loquat Cultivar with Large Fruit 'Huajin Wuhe No. 1'. *Acta Hortic. Sin*. 46, 2763–2764. doi: 10.16420/j.issn.0513-353x.2019-0241

Dang, J., Guo, Q., Xiang, S., He, Q., Sun, H., Wu, D., et al. (2019b). A New Seedless Loquat Cultivar 'Huayu Wuhe No. 1' with Large Fruit. *Acta Hortic. Sin*. 2776–2767. doi: 10.16420/j.issn.0513-353x.2019-0242

Dawe, R. K. (1998). Meiotic chromosome organization and segregation in plants. *Annu. Rev. Plant Biol*. 49, 371–395. doi: 10.1146/annurev.arplant.49.1.371

Dawe, R. K., Sedat, J. W., Agard, D. A., and Cande, W. Z. (1994). Meiotic chromosome pairing in maize is associated with a novel chromatin organization. *Cell*. 76, 901–912. doi: 10.1016/0092-8674(94)90364-6

De Muyt, A., Pereira, L., Vezon, D., Chelysheva, L., Gendrot, G., Chambon, A., et al. (2009). A high throughput genetic screen identifies new early meiotic recombination functions in *Arabidopsis thaliana*. *PLOS Genetics*. 5, e1000654. doi: 10.1371/journal.pgen.1000654

De Storme, N., and Geelen, D. (2011). The arabidopsis mutant jason produces unreduced first division restitution male gametes through a parallel/fused spindle mechanism in meiosis II. *Plant Physiol*. 155, 1403–1415. doi: 10.1104/pp.110.170415

De Vries, F. A. T., de Boer, E., van den Bosch, M., Baarends, W. M., Ooms, M., Yuan, L., et al. (2005). Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev*. 19, 1376–1389. doi: 10.1101/gad.329705

Deng, Q. (2009a). Studies on Embryological Mechanism of Seed Degeneration and Genetic Diversity of Seedlings from Middle-Degenerated Seeds in Loquat (*Eriobotrya japonica* Lindl.). [dissertation/Doctoral thesis]. Sichuan: Sichuan Agricultural University.

Deng, X., Guo, W., and Sun, X. (1996). Research progress on the type selection of seedless citrus in China literature review. *Acta Hortic. Sin*. 23, 235–240.

Deng, Y. (2009b). Research on the Nucle-free induction mechanism of GA3 in Eriobotrya japonica. [Dissertation/doctoral thesis]. Chongqing: Southwest University.

Deng, Z., and Wang, T. (2007). OsDMC1 is required for homologous pairing in *Oryza sativa*. *Plant Mol. Biol*. 65, 31–42. doi: 10.1007/s11103-007-9195-2

d'Erfurth, I., Jolivet, S., Froger, N., Catrice, O., Novatchkova, M., Simon, M., et al. (2008). Mutations in AtPS1 (*Arabidopsis thaliana* Parallel Spindle I) Lead to the Production of Diploid Pollen Grains. *PLOS Genetics*. 4:e1000274. doi: 10.1371/journal.pgen.1000274

Dong, M., Li, J., Zhou, D., Yue, J., and Gao, J. (2013). Advances in citrus cultivars breeding. Chinese Fruit Trees. p, 73–78.

Edlinger, B., and Schlogelhofer, P. (2011). Have a break: determinants of meiotic DNA double strand break (DSB) formation and processing in plants. *J. Exp. Bot*. 62, 1545–1563. doi: 10.1093/jxb/erq421

Fan, P., Yang, M., Zhang, Y., and Li, S. (2004). Early-ripening Seedless Grape 'Jingzaojing'. *Acta Hortic. Sin*. 31, 415. doi: 10.16420/j.issn.0513-353x.2004.03.036

Ferdous, M., Higgins, J. D., Osman, K., Lambing, C., Roitinger, E., Mechtler, K., et al. (2012). Inter-homolog crossing-over and synapsis in *Arabidopsis* meiosis are dependent on the chromosome axis protein AtASY3. *PLOS Genet*. 8, e1002507. doi: 10.1371/journal.pgen.1002507

Gambetta, G., Gravina, A., Fasiolo, C., Fornero, C., Galiger, S., Inzaurralde, C., et al. (2013). Self-incompatibility, parthenocarpy and reduction of seed presence in 'afourer' mandarin. *Sci. Hortic*. 164, 183–188. doi: 10.1016/j.scienta.2013.09.002

Garmendia, A., Beltrán, R., Zornoza, C., Breijo, F., Reig, J., Bayona, I., et al. (2019). Insect repellent and chemical agronomic treatments to reduce seed number in 'afourer' mandarin. effect on yield and fruit diameter. *Sci. Hortic*. 246, 437–447. doi: 10.1016/j.scienta.2018.11.025

Gillies, C. B., and Moens, P. B. (1984). The synaptonemal complex in higher plants. *Crit. Rev. Plant Sci*. 2, 81–116. doi: 10.1080/07352688409382191

Grandont, L., Cunado, N., Coriton, O., Huteau, V., Eber, F., Chevre, A. M., et al. (2014). Homoeologous chromosome sorting and progression of meiotic recombination in brassica napus: ploidy does matter! *Plant Cell*. 26, 1448–1463. doi: 10.1105/tpc.114.122788

Gu, J., and Zhang, S. (1990). Production of seedless loquat. *Foreign agriculture*. 3, 31–33.

Guo, J., Sun, R., Song, J., and Zhang, S. (2001). Microsporogenesis of Several-Male 2 Sterile Lines in Brassica rapa L. ssp *pekinensis*. *Acta Hortic. Sin*. 28, 409–414. doi: 10.3321/j.issn:0513-353X.2001.05.005

Guo, Q., Li, X., Xiang, S., He, Q., Sun, H., Wu, D., et al. (2016). A new white pulp seedless loquat cultivar 'Wuhe Guoyu'. *Acta Hortic. Sin*. 43, 2717–2718. doi: 10.16420/j.issn.0513-353x.2016-0445

Higgins, J. D., Sanchez-Moran, E., Armstrong, S. J., Jones, G. H., and Franklin, F. C. H. (2005). The *Arabidopsis* synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev*. 19, 2488–2500. doi: 10.1101/gad.354705

Hollingsworth, N. M., Goetsch, L., and Byers, B. (1990). The hop1 gene encodes a meiosis-specific component of yeast chromosomes. *Cell*. 61, 73–84. doi: 10.1016/0092-8674(90)90216-2

Hong, L., Tang, D., Zhu, K., Wang, K., Li, M., and Cheng, Z. (2012). Somatic and reproductive cell development in rice anther is regulated by a putative glutaredoxin. *Plant Cell*. 24, 577–588. doi: 10.1105/tpc.111.093740

Hoyt, M. A., and Geiser, J. R. (1996). Genetic analysis of the mitotic spindle. *Annu. Rev. Genet*. 30:7–33. doi: 10.1146/annurev.genet.30.1.7

Hu, S. (1982). *Angiosperms Embryology*. Beijing: Higher Education Press. p, 54.

Hu, S. (1993). Experimental methods in plant embryology (I) determination of pollen viability. *Chin. J. Bot*. 10, 60–62.

Huang, J., and Xu, X. (1980). Loquat new varieties 'TaiCheng No.4'. *China Fruits*. 1, 35–37.

Huang, J., Xu, X., and Chen, X. (1984). Cultivation of tetraploid loquat 'Min NO.3'. *Chinese Fruit Trees*. p, 27–30.

Ji, J., Tang, D., Shen, Y., Xue, Z., Wang, H., Shi, W. Q., et al. (2016). P31comet, a member of the synaptonemal complex, participates in meiotic DSB formation in rice. *Proc. Natl. Acad. Sci. U.S.A*. 113, 10577–10582. doi: 10.1073/pnas.1607334113

Jiang, F., Huang, A., Chen, Z., Deng, C., Chen, X., Chen, X., et al. (2009). Study on seed traits in Loquat (Eriobotrya Japonica Lindl.) germplasm resources. *Fujian. Hortic*. 4, 19–24. doi: 10.3969/j.issn.1004-6089.2009.04.005

Jones, G. H. (1984). The control of chiasma distribution. *Symp. Soc. Exp. Biol*. 38, 293–320.

Kawanabe, T., Ariizumi, T., Kawai-Yamada, M., Uchimiya, H., and Toriyama, K. (2006). Abolition of the tapetum suicide program ruins microsporogenesis. *Plant Cell Physiol*. 47, 784–787. doi: 10.1093/pcp/pcj039

Kikuchi, S., Iwasuna, M., Kobori, A., Tsutaki, Y., Yoshida, A., Murota, Y., et al. (2014). Seed formation in triploid loquat (Eriobotrya japonica) through cross-hybridization with pollen of diploid cultivars. *Breed*. 64,176–182. doi: 10.1270/jsbbs.64.176

Lange, T. D. (1998). Telomeres and senescence: ending the debate. *Science*. 279, 334–335. doi: 10.1126/science.279.5349.334

Ledbetter, C. A., and Ramming, D. W. (1989). "Seedlessness in Grapes", in *Horticultural Reviews II,* ed. J. Janick (Portland: Timber Press). p, 159–184. doi: 10.1002/9781118060841.ch5

Levan, A., Fredga, K., and Sandberg, A. A. (1964). Nomenclature for centromeric position of chromosomes. *Hereditas*. 52, 201–220. doi: 10.1111/j.1601-5223.1964.tb01953.x

Li, M., and Chen, R. (1985). A suggestion on the standardization of karyotype analysis in plants. *J. Wuhan Bot. Res*. 3, 297–302.

Li, M., Lu, C., Liu, X., Wang, X., and Zheng, X. (2016). Research progress of seedless and stenospermocarpic mechanism in litchi. *J. Trop. Crop Sci*. 37, 7. doi: 10.3969/j.issn.1000-2561.2016.05.030

Li, N., Zhang, D., Liu, H., Yin, C., and Zhang, D. (2006). The rice tapetum degeneration retardation gene is required for tapetum degradation and anther development. *Plant Cell*. 18, 2999–3014. doi: 10.1105/tpc.106.044107

Li, S., and Wang, Y. (2019). Advances in seedless gene researches and seedless breeding in grapevine. *Acta Hortic. Sin*. 9, 16. doi: 10.16420/j.issn.0513-353x.2019-0035

Li, Z. (1987). *Plant Section Technology*. Beijing: Science Press.

Li, Z., Luo, Q., and Wang, Y. (2019). Breeding seedless grapevine *via* embryo rescue and marker-assisted selection in hybrid progenies. *J. Fruit Sci*. 36, 12. doi: 10.13925/j.cnki.gsxb.20180098

Liang, G., and Li, X. (1991). Improved smear method of plant meiosis on citrus. *South China Fruits*. 2, 40–41.

Liang, S., Dang, J., Liang, G., and Guo, Q. (2018). Meiosis Observation and Fertility Analysis in Natural Tetraploid Loquat of 'B431'. *Acta Hortic. Sin*. 45, 1895–1904. doi: 10.16420/j.issn.0513-353x.2018-0222

Lin, S. (2008). "Loquat," in: *Encyclopedia of Fruits and Nuts* (Wallingford: CABI). p, 643–651.

Lin, S., Sharpe, R. H., and Janick, J. (1999). Loquat: Botany and horticulture. *Hortic. Rev*. 23, 233–276. doi: 10.1002/9780470650752.ch5

Loidl, J. (1990). The initiation of meiotic chromosome pairing: the cytological view. *Genome*. 33, 759. doi: 10.1139/g90-115

Lui, D. Y., and Colaiacovo, M. P. (2013). Meiotic development in Caenorhabditis elegans. *Adv. Exp. Med. Biol*. 757, 133–170. doi: 10.1007/978-1-4614-4015-4_6

Maria, L., Janick, J., Lin, S., Wang, W., Liang, G., and Wang, W. (2013). "Breeding Loquat", in *Plant Breeding Reviews*, ed. by J. Janick, (New Jersey, NJ: John Wiley and Sons, Inc.). p, 259–296. doi: 10.1002/9781118497869.ch5

Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N., and Grelon. M. (2015). The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol*. 66, 297–327. doi: 10.1146/annurev-arplant-050213-035923

Miao, C., Tang, D., Zhang, H., Wang, M., Li, Y., Tang, S., et al. (2013). Central region component1, a novel synaptonemal complex component, is essential for meiotic recombination initiation in rice. *Plant Cell*. 25, 2998–3009. doi: 10.1105/tpc.113.113175

Moens, P. B. (1969). The fine structure of meiotic chromosome polarization and pairing in locusta migratoria spermatocytes. *Chromosoma*. 28, 1–25. doi: 10.1007/BF00325986

Nimmo, E. R., Pidoux, A. L., Perry, P. E., and Allshire, R. C. (1998). Defective meiosis in telomere-silencing mutants of schizosaccharomyces pombe. *Nature*. 392, 825–828. doi: 10.1038/33941

Page, S. L., and Hawley, R. S. (2001). c (3)G encodes a *Drosophila* synaptonemal complex protein. *Gene Dev*. 15, 3130–3143. doi: 10.1101/gad.935001

Pearson, H. M. (1932). Parthenocarpy and seedlessness in *Vitis vinifera*. *Science*. 76, 594–594. doi: 10.1126/science.76.1982.594.a

Peoples-Holst, T. L., and Burgess, S. M. (2005). Multiple branches of the meiotic recombination pathway contribute independently to homolog pairing and stable juxtaposition during meiosis in budding yeast. *Gene Dev*. 19, 863–874. doi: 10.1101/gad.1293605

Qiu, W., Zhu, A., Wang, Y., Chai, L., Ge, X., Deng, X., et al. (2012). Comparative transcript profiling of gene expression between seedless ponkan mandarin and its seedy wild type during floral organ development by suppression subtractive hybridization and cdna microarray. *BMC Genomics*. 13, 397. doi: 10.1186/1471-2164-13-397

Rédei, G. P. (2008). Encyclopedia of genetics, genomics, proteomics, and informatics. *Springer Science and Business Media*. Volume 1 A-L, p. 1773. doi: 10.1007/978-1-4020-6754-9

Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Bourbon, H. M., et al. (2016). The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science*. 351, 943–949. doi: 10.1126/science.aad5309

Rocha, M., Chiavegatto, R. B., Damasceno, A. G., Rocha, L. C., and Techio, V. H. (2019). Comparative meiosis and cytogenomic analysis in euploid and aneuploid hybrids of *Urochloa* P. beauv. *Chromosome Res.* 27, 333–344. doi: 10.1007/s10577-019-09616-y

Royo, C., Carbonell-Bejerano, P., Torres-Pérez, R., Nebish, A., Martínez, Ó., Rey, M., et al. (2016). Developmental, transcriptome, and genetic alterations associated with parthenocarpy in the grapevine seedless somatic variant Corinto bianco. *J. Exp. Bot*. 67:259–273. doi: 10.1093/jxb/erv452

Sanchez-Moran, E., Osman, K., Higgins, J. D., Pradillo, M., Cuñado, N., Jones,. G. H., et al. (2008). ASY1 coordinates early events in the plant meiotic recombination pathway. *Cytogenet. Genome Res*. 120, 302–312. doi: 10.1159/000121079

Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E., and Schneitz, K. (1999). Molecular analysis of nozzle, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana. in: *Proceedings of the National Academy of Sciences of the United States of America* (Washington, DC). 96, 11664–11664. doi: 10.1073/pnas.96.20.11664

Sepsi, A., and Schwarzacher, T. (2020). Chromosome-nuclear envelope tethering-a process that orchestrates homologue pairing during plant meiosis? *J. Cell Sci.* 133:jcs243667. doi: 10.1242/jcs.243667

Siddiqi, I., Ganesh, G., Grossniklaus, U., and Subbiah, V. (2000). The dyad gene is required for progression through female meiosis in arabidopsis. *Development*. 127, 197–207. doi: 10.1242/dev.127.1.197

Smith, A. V., and Roeder, G. S. (1997). The Yeast Red1 Protein Localizes to the Cores of Meiotic Chromosomes. *J. Cell Biol*. 136, 957–967. doi: 10.1083/jcb.136.5.957

Smith, M. B., Palmer, R. G., and Horner, H. T. (2002). Microscopy of a cytoplasmic male-sterile soybean from an interspecific cross between Glycine max and *G. soja* (Leguminosae). *Am. J. Bot*. 89, 417–417. doi: 10.3732/ajb.89.3.417

Stebbins, G. L. (1971). Chromosomal evolution in higher plants. London: Arnold. p, 216.

Suzuki, K., Takeda, H., Tsukaguchi, T., and Egawa, Y. (2001). Ultrastructural study on degeneration of tapetum in anther of snap bean (*Phaseolus vulgaris L.*) under heat stress. *Sex. Plant Reprod*. 13, 293–299. doi: 10.1007/s004970100071

Sym, M., Engebrecht, J., and Roeder, G. S. (1993). Zip1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell*. 72, 365–378. doi: 10.1016/0092-8674(93)90114-6

Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J., and Stahl, F. W. (1983). The double-strand-break repair model for recombination. *Cell*. 33, 25–35. doi: 10.1016/0092-8674(83)90331-8

Tang, X. (2001). Study on seedless or less seedless mechanism and seedless type cultivation of persimmon. [dissertation/doctoral thesis]. Wuhan: Huazhong Agricultural University.

Tian, S., Ge, J., Ai, G., Jiang, J., Liu, Q., Chen, X., et al. (2021). A 2.09 Mb fragment translocation on chromosome 6 causes abnormalities during meiosis and leads to less seed watermelon. *Hortic. Res*. 8:256. doi: 10.1038/s41438-021-00687-9

Trelles-Sticken, E., and Scherthan, D. H. (2000). Meiotic telomere protein ndj1p is required for meiosis-specific telomere distribution, bouquet formation and efficient homologue pairing. *J. Cell Biol*. 151, 95–106. doi: 10.1083/jcb.151.1.95

Varas, J., Graumann, K., Osman, K., Pradillo, M., Evans, D. E., Santos, J. L., et al. (2015). Absence of sun1 and sun2 proteins in *Arabidopsis thaliana* leads to a delay in meiotic progression and defects in synapsis and recombination. *Plant J*. 81, 329–346. doi: 10.1111/tpj.12730

Wang, H., Dang, J., Wu, D., Xie, Z., Yan, S., Luo, J., et al. (2021). Genotyping of polyploid plants using quantitative PCR: application in the breeding of white-fleshed triploid loquats (Eriobotrya japonica). *Plant Methods*. 17, 93. doi: 10.1186/s13007-021-00792-9

Wang, M., Wang, K., Tang, D., Wei, C., Li, M., Shen, Y., et al. (2010). The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice. *Plant Cell*. 22, 417–430. doi: 10.1105/tpc.109.070789

Xiao, J., Tan, J., Liu, H., Chen, L., Ye, W., and Cheng, W. (2007). Studies on the seedless mechanism of 'Lipeng No.2' ponkan (Citrus reticulata). *J. Fruit Sci*. 4, 421–426. doi: 10.13925/j.cnki.gsxb.2007.04.003

Xu, F., Zhang,. X., Shi,. C., Wang,. X., Xi,. X., Jiang,. S., et al. (2016). A study on the influence factors of single-seeded formation in loquat. *J. Fruit Sci*. 33, 676–685. doi: 10.13925/j.cnki.gsxb.20160099

Xu, X., Si, J., Xie, J., Lan, Y., Zeng, Y., Jiang, W., et al. (2006). Ougan seedless, a new mandarin cultivar. *J. Fruit Trees*. 5, 781–782. doi: 10.13925/j.cnki.gsxb.2006.05.031

Yang, F., Fuente, R., Leu, N. A., Baumann, C., McLaughlin, K. J., and Wang, P. (2006). Mouse SYCP2 is required for synaptonemal complex assembly and chromosomal synapsis during male meiosis. *J. Cell Biology*. 173, 497–507. doi: 10.1083/jcb.200603063

Yang, W., Ye, D., Xu, J., and Sundaresan, V. (1999). The SPOROCYTELESS gene of Arabidopsis is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes Dev*. 13, 2108–2117. doi: 10.1101/gad.13.16.2108

Yang, Y. (2021). Analysis of female fertility of triploid loquat (Eriobotrya japonica) Q24 and preliminary study on genome characteristics. [dissertation/master's thesis]. Chongqing: Southwest University.

Ye, Z., Zeng, T., Xu, J., Luo, Z., Hu, G., Zhang, Z., et al. (2006). Breeding of seedless tangerine (shiyutangerine). *J. Fruit Sci*. 23,149–150+F0002. doi: 10.1016/j.compscitech.2006.07.013

Yu, H., Wang, M., Tang, D., Wang, K., Chen, F., Gong, Z., et al. (2010). OsSPO11-1 is essential for both homologous chromosome pairing and crossover formation in rice. *Chromosoma*. 119, 625–636. doi: 10.1007/s00412-010-0284-7

Zhang, G., Kang, L., Gao, Z., Zhu, S., and Gao, K. (1999). Effects of Ga and CPPU on quality of loquat seedless fruit. *J. Fruit Sci*. 16, 55–59.

Zhang, S., Huang, G., Ding, F., He, X., and Pan, J. (2012). Mechanism of seedlessness in a new lemon cultivar 'Xiangshui' [Citrus limon (L.) Burm. F.]. *Sex. Plant Reprod*. 25:337–345. doi: 10.1007/s00497-012-0201-8

Zhang, W., Sun, Y., Timofejeva, L., Chen, C., and Ma, H. (2006). Regulation of Arabidopsis tapetum development and function by Dysfunctional Tapetum1 (DYT1) encoding a putative bHLH transcription factor. *Development*. 133, 3085–3095. doi: 10.1242/dev.02463

Zhang, X., Liao, M., He, J., Liu, C., Ma, Q., Yang, D., et al. (2015). Study on Stigma Receptivity and Pollen Tube growth of 'Chuanzao loquat' in the First Florescence. *Acta Bot. Sin*. 1349–1355. doi: 10.7606/j.issn.1000-4025.2015.07.1349

Zhang, X., Liao, M., Wang, Y., Li, X., and Tao, L. (2014). Research on the Difference in the Pollen Tube Growth of a Low Seediness Line of Loquat. *Acta Bot. Sin*. 34, 26–31. doi: 10.7606/j.issn.1000-4025.2014.01.0026

Zilli, A. L., Brugnoli, E. A., Marcón, F., Billa, M. B., Rios, E. F., Martínez, E. J., et al. (2015). Heterosis and expressivity of apospory in tetraploid bahiagrass hybrids. *Crop Sci*. 55, 1189–1201. doi: 10.2135/cropsci2014.10.0685

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fpls.2022.884957

Metabolomic and Transcriptomic Profiling Uncover the Underlying Mechanism of Color Differentiation in *Scutellaria baicalensis* Georgi. Flowers

*Defu Wang†, Jiangran Wang†, Yufen Wang, Dongzuo Yao and Yanbing Niu**

College of Life Sciences, Shanxi Agricultural University, Jinzhong, China

OPEN ACCESS

Edited by:
Yuexue Liu, Shenyang Agricultural University, China

Reviewed by:
Yunpeng Cao, Chinese Academy of Sciences (CAS), China
Kang-Di Hu, Hefei University of Technology, China
Shiqing Gao, Beijing Academy of Agricultural and Forestry Sciences, China

***Correspondence:**
Yanbing Niu
niuyanbingbest@163.com

†These authors have contributed equally to this work and share first authorship

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

Received: *27 February 2022*
Accepted: *10 May 2022*
Published: *09 June 2022*

Citation:
Wang D, Wang J, Wang Y, Yao D and Niu Y (2022) Metabolomic and Transcriptomic Profiling Uncover the Underlying Mechanism of Color Differentiation in Scutellaria baicalensis Georgi. Flowers. Front. Plant Sci. 13:884957. doi: 10.3389/fpls.2022.884957

Scutellaria baicalensis Georgi. (Chinese skullcap or Huang-qin) is an extremely crucial medicinal plant in the Labiate family, and the color of its flowers naturally appears purple. However, during the long-term cultivation of *S. baicalensis*, very few plants of *S. baicalensis* also present white and purple-red flower colors under the same ecological conditions. However, the complex metabolic and transcriptional networks underlying color formation in white, purple-red, and purple flowers of *S. baicalensis* remain largely unclarified. To gain an insight into this issue, we conducted transcriptome and metabolomic profiling to elucidate the anthocyanin synthesis metabolic pathway in the flowers of *S. baicalensis*, and to identify the differentially expressed candidate genes potentially involved in the biosynthesis of anthocyanins. The results showed that 15 anthocyanins were identified, among which cyanidin 3-rutinoside and delphin chloride were the primary anthocyanins, and accumulation was significantly related to the flower color changes of *S. baicalensis*. Furthermore, the down-regulation of *SbDFR* (*Sb02g31040*) reduced the anthocyanin levels in the flowers of *S. baicalensis*. The differential expression of the *Sb3GT* (*Sb07g04780* and *Sb01g72290*) gene in purple and purple-red flowers affected anthocyanin accumulation, suggesting that anthocyanin levels were closely associated with the expression of *SbDFR* and *Sb3GT*, which play important roles in regulating the anthocyanin biosynthesis process of *S. baicalensis* flowers. Transcriptomic analysis revealed that transcription factors WRKY, bHLH, and NAC were also highly correlated with anthocyanin accumulation, especially for NAC35, which positively regulated *SbDFR* (*Sb02g31040*) gene expression and modulated anthocyanin biosynthesis in flower color variation of *S. baicalensis*. Overall, this study presents the first experimental evidence for the metabolomic and transcriptomic profiles of *S. baicalensis* in response to flower coloration, which provides a foundation for dynamic metabolic engineering and plant breeding, and to understand floral evolution in *S. baicalensis* plants.

Keywords: anthocyanin, flowers, *S. baicalensis*, transcriptome, metabolic profiling

INTRODUCTION

Scutellaria baicalensis Georgi. (Chinese skullcap or Huang-qin) is an erect, perennial herb belongs to the Labiate family, which has been cultivated for its therapeutic properties in China for 2000 years (Shang et al., 2010; Zhao et al., 2016a,b). The dried root of *S. baicalensis* is rich in flavonoids and used extensively as a traditional medicine for treating fever and lung and liver complaints (Li, 2012). Modern pharmacological research shows that the bioactivity of root flavonoids from *Scutellaria* also has antibacterial, antiviral, antioxidant, anticancer, hepatoprotective, and neuroprotective properties (Baumann et al., 2008; Gao et al., 2011; Yang et al., 2012; Nayak et al., 2014).

The flower color of most *S. baicalensis* usually appears purple in nature and in the field, while our research group found that some plants of *S. baicalensis* have white and purple-red flowers under the same ecological conditions. Flower color is affected by various external and internal factors, but anthocyanin type and content are some of the most significant factors that may influence flower color. Anthocyanins, which are natural colorants belonging to an important subgroup of flavonoids, cannot only make plant tissues and organs show different colors but also help plants resist biological and abiotic stress, including protecting plants from pathogenic bacteria infection, resisting ultraviolet radiation, and removing excess active oxygen (Chen et al., 2012; Wang et al., 2013). Thus far, about 635 anthocyanins have been identified (He and Giusti, 2010). Cyanidin, delphinidin, pelargonidin, peonidin, petunidin, and malvidin are six common anthocyanins in the plant kingdom (Jaakola, 2013). Among them, delphinidin, malvidin, and petunidin are important coloring substances in many blue-purple plant organs; peonidin and cyanidin are the main pigments in purple-red plant organs, while pelargonin appears in brick red plant organs. These compounds have been reported to change the color of flowers from pink to blue violet (Tanaka et al., 2009; Peng et al., 2021). Anthocyanin biosynthesis is primarily associated with the pathway of phenylalanine metabolism (Dooner et al., 1991; Silva et al., 2007). The key enzymes involved in the earlier stages of anthocyanin synthesis are chalcone isomerase (CHI), chalcone synthase (CHS), flavonoid 3′-hydroxylase and flavanone 3-hydroxylase (F3H); while those in the later stages are anthocyanidin 3-glycosyltransferase, anthocyanidin synthase (ANS) and dihydroflavonol 4-reductase (DFR). At present, the genes responsible for anthocyanin biosynthesis have been found in *Ipomoea batatas* (Wang et al., 2013), *Ginkgo biloba* (Cheng et al., 2013), mulberry (Li D. et al., 2020), alfalfa (Duan et al., 2020), and other plants. It has been reported that color changes may occur in fruit, leaves and flowers, and the overexpression of enzymes-encoding genes and single-gene mutations can affect anthocyanin accumulation and color variation (Han et al., 2012; Xie et al., 2012; Li et al., 2013).

Apart from the above-mentioned biosynthetic genes, many transcription factors (TFs) are involved in the modulation of anthocyanin synthesis pathway and color changes (Saito et al., 2013). The roles of helix-loop-helix (bHLH), R2R3-MYB and WD40 have been widely investigated in recent years, which can form a MYB-bHLH-WD40 (MBW) complex to directly regulate the expression of genes related to anthocyanin synthesis pathway, thus promoting the biosynthesis of anthocyanins (Gonzalez et al., 2008; Suzuki et al., 2016; Bai et al., 2019; Chen et al., 2019). In addition, the formation of anthocyanins also depends on acylation, glycosylation, hydroxylation and methoxylation to improve stability, and some TFs play a crucial role in the transport of anthocyanins in plants (Lou et al., 2014; Wu et al., 2016).

Recently, a combination of high-throughput approaches has been proposed to assess color changes. The integrated analysis of metabolomic and transcriptomic data have revealed the altered secondary metabolites and differentially expressed genes (DEGs) in flowers and fruit (Wang et al., 2017; Dong et al., 2019; Meng et al., 2019; Zhuang et al., 2019; Li H. et al., 2020; Zhang et al., 2020), thus providing a global view of plant color development. The present study aimed to explore the transcriptional and metabolic profiles of three different flower colors in *S. baicalensis*, and the DEGs that responded to anthocyanin biosynthesis, three cultivars of *S. baicalensis* with different flower colors (SbP, SbPR, and SbW) were employed. Transcriptional, metabolic, and integrative analyses were conducted on three *S. baicalensis* plants with different flower colors. The findings provide important insights into the underlying mechanism of anthocyanin accumulation and the anthocyanin metabolic pathway in *S. baicalensis* flowers, which lay a substantive foundation for future molecular and functional studies in the creation of new *S. baicalensis* germplasm with different petal colors to enrich agricultural landscape planning and design.

MATERIALS AND METHODS

Plant Materials

Three kinds of *S. baicalensis* with different petal colors were cultivated at the *S. baicalensis* germplasm resource center at Shanxi Agricultural University, China, and were termed 'purple petals' (SbP), 'purple-red petals' (SbPR), and 'white petals' (SbW). Fresh purple, purple-red, and white petals were harvested from healthy *S.* baicalensis at the same development stage in July 2020. All materials were snap-frozen in liquid nitrogen and stored at −80°C for metabolite extraction and RNA sequencing. All experiments in this study were conducted using three replicates.

Measurement of Relative Anthocyanin Content

After the plants of *Scutellaria* blossomed, 0.1 g of petal of the corresponding groups were collected, ground with 1 mL methanol (0.1% HCl), rinsed for two times, and placed in a 10-mL centrifuge tube. The samples were eluted with methanol (0.1% HCl) to a final volume of 5 mL. The tissue homogenate was oscillated for 30 s, followed by centrifugation (12,000 g, 10–15 min, 4°C). The supernatant extract was filtered with a 0.22-μm inlet filter, and the optical density of the supernatants was detected using an ultraviolet spectrophotometer at 530 nm. The relative levels of anthocyanins were calculated as follows: $Q = V \times A_{530}/M$

(units/g FW), where V and M represent the volume of the solution and the weight of the samples. The concentration of anthocyanin at OD_{530} = 0.1 was regarded as 1 unit to calculate the relative content of anthocyanin in the samples. Methanol (0.1% HCl) was employed as a blank control (Jiang et al., 2020).

Sample Preparation for Metabolomic Analysis

The vacuum freeze-dried *S. baicalensis* specimens were ground to powder using a mixer mill (MM 400, Verder Retsch, Shanghai, China) at 30 Hz for 90 s. During extraction, 100 mg of powder was weighed and dissolved in 1 mL aqueous methanol (70%). Subsequently, the mixture was incubated overnight at 4°C in a refrigerator. During incubation, the samples were swirled three times at 10-min intervals (10 s, 40 Hz) to improve the extraction rate. After centrifugation (10,000 g, 10 min), the extract was absorbed using a CNWBOND Carbon-GCB SPE Cartridge (250 mg, 3 mL; ANPEL, Shanghai, China)[1] and filtrated (SCAA-104, 0.22-μm pore size; ANPEL, Shanghai, China)(see text footnote 1) prior to ultra-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) analysis.

Ultra-Performance Liquid Chromatography Conditions and ESI-Q TRAP-MS/MS

The extracted samples were subjected to LC-ESI-MS/MS analysis (UPLC, Shim-pack UFLC SHIMADZU CBM30A[2], MS/MS, Applied Biosystems 4500 Q TRAP[3]) (Chen et al., 2013). The following conditions were used: ACQUITY UPLC HSS T3 (C18) columns (Waters; 2.1 mm × 100 mm, 1.8 μm). The mobile phase was composed of solvent A (ultra-pure water with 0.1% formic acid) and solvent B (acetonitrile). The following gradient program was applied: 0 min V(A)/V(B) (100:0), 11.0 min V(A)/V(B) (5:95), 12.0 min V(A)/V(B) (5:95), 12.1 min V(A)/V(B) (95:5), and 15.0 min V(A)/V(B) (95:5). The injection volume, column oven temperature and flow rate were 5 μL, 40°C, and 0.4 mL/min, respectively. Mass spectrometric detection was performed in electrospray ionization (ESI) positive mode. The effluents were connected alternatively to an ESI-QTRAP-MS/MS. The standards were detected with a gas temperature of 550°C, capillary voltage of 5500 V, and nebulizer pressure of 25 psi.

Identification and Quantification of Metabolites

The metabolites were identified and quantified by the Wuhan MetWare Biotechnology Co., Ltd. (Wuhan, China). The scheduled multiple reaction monitoring (MRM) method previously described by Fraga et al. (2010) was used. Metabolites were identified using the MetWare MWDB database and public database of MassBank[4], KNAPSAcK[5], HMDB[6], MoTo DB[7], and METLIN[8] (Wishart et al., 2013; Zhu et al., 2013). The filtering conditions for the differentially accumulated metabolites (DAMs) were as follows: absolute log2 (fold-change) ≥ 1, variable importance in projection (VIP) ≥ 1 and p-value < 0.05. To determine the specific DAMs, R software[9] was employed to conduct principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA). Pathway enrichment analysis on differential metabolites was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Zhuang et al., 2019).

RNA Extraction, Illumina Sequencing, and Annotation

TRIzol reagent (Invitrogen, Carlsbad, CA, United States) was applied to extract total RNA from different flowers by following the kit's protocol. The RNA concentration and purity of each sample were determined by NanoDrop ND-1000 (NanoDrop Technologies, Wilmington, DE, United States). Meanwhile, the RNA integrity was assessed by the Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, United States). mRNA was enriched from the total RNA using magnetic beads with oligo (dT) primers, and cDNA was synthesized by SuperScript™ II Reverse Transcriptase (Invitrogen, cat. 1896649, Carlsbad, CA, United States) and linking the sequencing adapter to both ends. The libraries were then sequenced on an Illumina Novaseq™ 6000 platform (LC-Bio Technology Co., Ltd., Hangzhou, China). Clean reads were extracted with base-pair qualities of Q ≥ 20 via custom Perl scripts, followed by mapping on the *S. baicalensis* genome using the HisAT2 software with default settings.

Screening of Differentially Expressed Genes

The differentially expressed genes (DEGs) were selected with the thresholds of log2 (fold-change) ≥ 1 and false discovery rate (FDR) < 0.05 using the edge package in R (Robinson et al., 2010). Gene ontology (GO) enrichment analysis was implemented using the topGO package in R (corrected p-value < 0.05). The KEGG database[10] was used to conduct pathway analysis (Kanehisa et al., 2008).

Integrative Analysis of Transcriptomic and Metabolomic Profiles

The DEGs from anthocyanins and TFs responsible for the biosynthesis of anthocyanins were correlated with 15 DAMs of anthocyanins. Integrative analysis was conducted by calculating the FPKM values of genes and metabolites, and the screening

[1]http://www.anpel.com.cn/

[2]www.shimadzu.com.cn/

[3]www.appliedbiosystems.com.cn/

[4]http://www.massbank.jp/

[5]http://www.knapsackfamily.com/KNApSAcK/

[6]http://www.hmdb.ca/

[7]https://ngdc.cncb.ac.cn/databasecommons/database/id/3091

[8]http://metlin.scripps.edu/index.php

[9]www.r-project.org/

[10]http://www.genome.jp/kegg

criterion for DEGs was $|r| > 0.8$, and $|r| > 0.9$ and p-value < 0.05 for metabolites. Using an online tool (Lianchuan Cloud Biological Platform)[11] to reveal the interactive networks among DEGs and DAMs of anthocyanins. A heatmap was drawn based on DEGs and 15 anthocyanins using R software.

Quantitative Real-Time Polymerase Chain Reaction Assays

For use in gene expression validation of RNAseq, total RNA was isolated from *Scutellaria* flowers and analyzed by quantitative real-time polymerase chain reaction (qRT-PCR) protocols. Twelve genes associated with anthocyanin biosynthesis were randomly selected from RNAseq for qRT-PCR analysis, and the constitutively expressed *actin* was employed as a housekeeping gene. All specific primer pairs are listed in **Supplementary Table 1**. The 25-μL reaction mixture consisted of 1 μL first-strand cDNA, 20 μL 1 × SYBR Premix ExTaq (Takara), 1 μL forward primer (10 μM), 1 μL reverse primer (10 μM), and 2 μL highly pure water. The reaction was performed in an ABI 7500 fast (ABI, Fort Lauderdale, FL, United States). This experiment conducted in triplicate. The expression level of each target gene was determined by the $2^{-\Delta\Delta Ct}$ method using the *actin* gene as an internal standard (Schmittgen and Livak, 2008).

Sequence Alignment, Phylogenetic Analysis, and Three-Dimensional Structure Prediction of the Candidate gene

Multiple amino acid sequence alignment was conducted with DNAMAN software (Lynnon Corp., Canada). Phylogenetic analysis was performed using MEGA 7.1 by the maximum likelihood method with 1,000 bootstrap replicates. The 3D structures of the candidate gene were predicted using the ITASSER V3.0 and V5.0 server.[12] UCSF Chimera V1.9 program[13] was used to visualize the protein data bank files retrieved from this server.

Statistical Analyses

Statistical tests were conducted with Excel 2010 software (Microsoft Office, Redmond, WA, United States). All values are presented as mean ± standard deviation (SD). Least significant difference test was employed to compare the means between groups ($p < 0.05$).

RESULTS

Total Anthocyanin Content in Three Different Flowers of *Scutellaria baicalensis*

Although *S. baicalensis* was grown in the same ecological environment, the petal colors in different plants were obviously different, with SbW possessing white petals, SbPR possessing purple-red petals, and SbP possessing purple petals (**Figure 1A**). Anthocyanins are the major pigments in plants. In this work, the anthocyanin content in three different *S. baicalensis* flowers was measured. The relative anthocyanin levels of SbP and SbPR were 10.5 and 9.5 units/g of fresh weight, respectively, which were remarkably higher compared with SbW (3.5 units/g of fresh weight, **Figure 1B**).

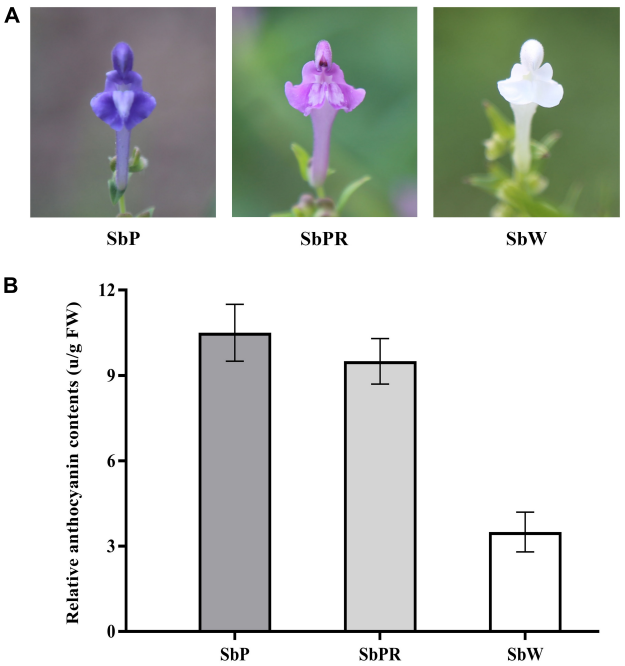


FIGURE 1 | *Scutellaria baicalensis* Georgi. flowers used in this study. **(A)** Phenotypes of *S. baicalensis* flowers. **(B)** Measurement of relative anthocyanin contents in flowers of SbW, SbPR, and SbP. SbW, white flower at the blooming stage; SbPR, purple-red flower at the blooming stage; SbP, purple flower at the blooming stage.

Metabolite Analysis on Three Different Flowers of *Scutellaria baicalensis*

Based on LC-MS widely targeted metabolome technology, qualitative and quantitative analyses of metabolites were conducted by MRM of triple quadrupole mass spectrometry. R software was used to construct a heatmap for metabolites. According to the differences in the accumulation of metabolites among different samples, hierarchical cluster analysis revealed that significant differences were found in metabolites among three different flowers of *S. baicalensis*, and three main clusters were obtained according to the relative differences of flavonoid accumulation patterns (**Figure 2A**). The flavonoids in clusters 1 accumulated at the highest levels in SbP followed by SbPR, and at the lowest levels in SbW. The flavonoids in cluster 2 were at the highest levels in SbPR followed by SbP, and were at the lowest levels in SbW. The flavonoids in cluster 3 were at the highest levels in SbW followed by SbPR, and were at the lowest levels in SbP. In total, 168 different flavonoid

[11]https://www.omicstudio.cn/tool/64

[12]http://zhanglab.ccmb.med.umich.edu/I-TASSER/

[13]https://www.cgl.ucsf.edu/chimera/

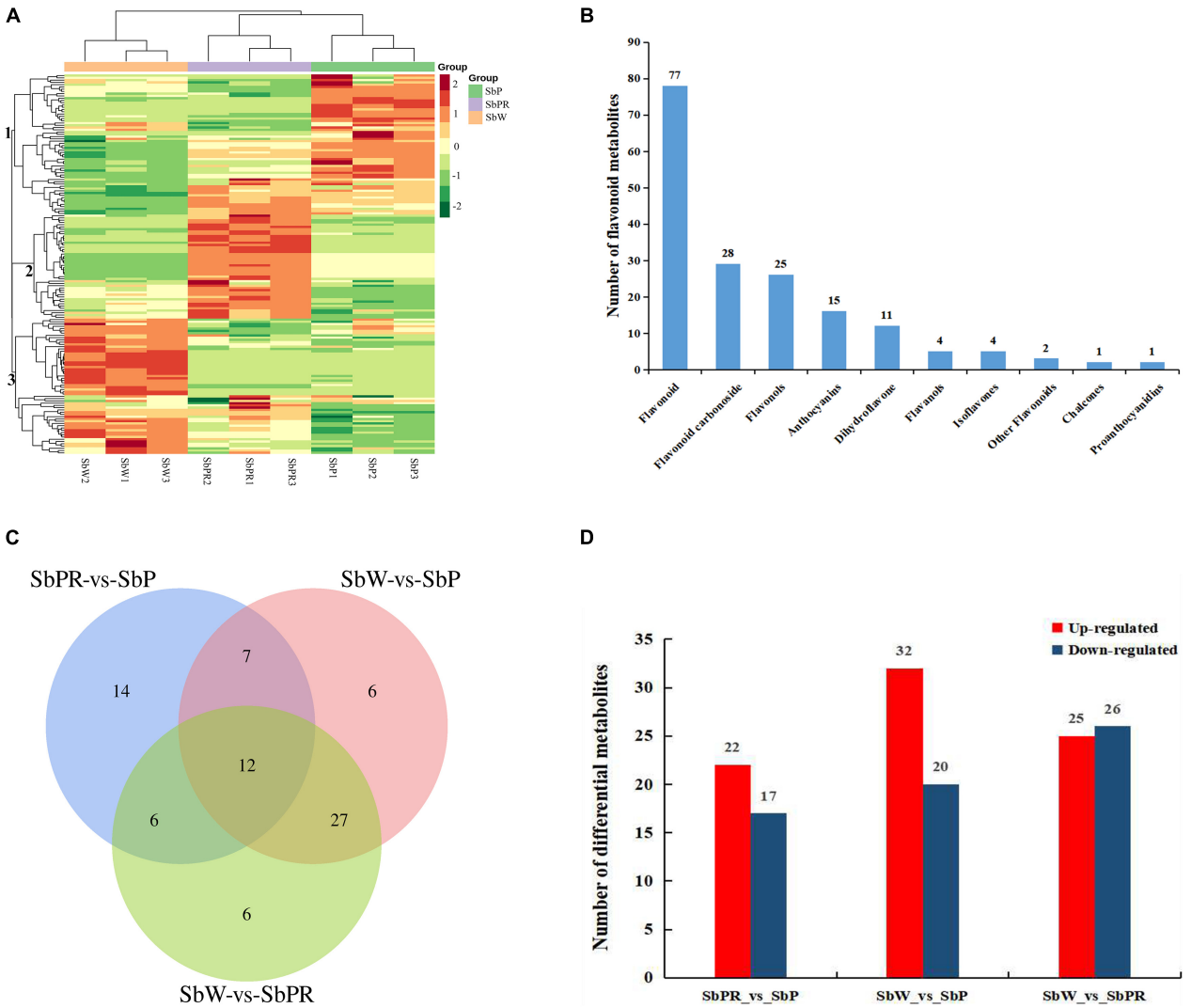


FIGURE 2 | Flavonoid metabolome profiles of the petals of white, purple-red, and purple flowers in *Scutellaria baicalensis*. **(A)** Heatmap clustering showing the correlation among white, purple-red, and purple flower samples based on 168 metabolite profiles. The color scale from red to green indicates the normalized metabolite levels using the row Z-score. **(B)** Classification of the 168 identified metabolites into 10 groups. **(C)** Venn diagram indicating the shared and specific number of the differential metabolites identified by each package. **(D)** Numbers of differential metabolites. SbW: white flower at the blooming stage; SbPR, purple-red flower at the blooming stage; SbP, purple flower at the blooming stage.

metabolites, such as 25 flavonols, 15 anthocyanins, 28 flavonoid carbonosides, 11 dihydroflavones, 4 flavanols, 4 isoflavones, 77 flavonoids, 1 chalcones, 1 proanthocyanidins, and 2 other flavonoids, were identified in the flowers of *S. baicalensis*. The details of all identified metabolites are summarized in **Figure 2B** and **Supplementary Table 2**.

To analyze the metabolic differences between SbW, SbPR, and SbP, screening of differential metabolites was conducted on all metabolites based on the following criteria: fold-change ≥ 1, p-value < 0.05 and VIP ≥ 1, which were graphed by Venn diagrams (**Figure 2C**). A total of 78 differential metabolites were identified among all samples and the relative contents of these differential metabolites were illustrated in a heatmap (**Supplementary Figure 1**). Among these differential metabolites, 12 differential anthocyanins were included, and their relative contents were described (**Supplementary Figure 2**). In addition, the overlapping metabolites were further identified. The results showed that 12 differential metabolites were shared among SbPR vs. SbP, SbW vs. SbP, and SbW vs. SbPR. These 12 differential metabolites included six flavonoids, two flavonols, one dihydroflavonols, and three anthocyanins. The three anthocyanins are classified as Cyanin chloride, Cyanidin 3-rutinoside and Delphin chloride. As demonstrated in **Figure 2D** and **Supplementary Table 3**, 39 differential metabolites (17 down-regulated and 22 up-regulated) were observed between SbPR and SbP, 52 differential metabolites (20 down-regulated and 32 up-regulated) between SbW and SbP, and 51 differential metabolites (26 down-regulated and 25 up-regulated) between

SbW and SbPR. Subsequently, the differential anthocyanins among these samples were obtained by analyzing the relative anthocyanin quantification results (**Supplementary Figure 3**). The results indicated that there were 5 (3 down-regulated and 2 up-regulated), 11 (0 down-regulated and 11 up-regulated), and 10 (1 down-regulated and 9 up-regulated) types of differential anthocyanins in SbPR vs. SbP, SbW vs. SbP, and SbW vs. SbPR, respectively, and these might be the key metabolites that influencing petal coloration in *S. baicalensis* (**Supplementary Figure 3**).

A total of 15 anthocyanins, such as cyanidin, delphinidin, pelargonidin, peonidin and malvidin, were identified in *S. baicalensis* flowers. Among them, the content of cyanidin 3-rutinoside in SbPR and SbP was significantly higher than that in SbW (105947- and 6828-fold, respectively), which suggested that it was an important color substance in *S. baicalensis*. The cyanidin 3-rutinoside and delphin chloride contents in SbPR were 645-fold higher and 53.8-fold lower, respectively, than those in SbP (**Table 1**). The significant difference in cyanidin 3-rutinoside and delphin chloride contents may be involved in the purple and purple-red flower colors of *S. baicalensis*.

Transcriptome Analysis of Three Different Flowers of *Scutellaria baicalensis*

To clarify the molecular mechanism of anthocyanin synthesis in three different flowers of *S. baicalensis*, transcriptomic analysis was conducted to identify DEGs in three different flowers. RNA-seq produced 40,138,441, 44,209,405, and 41,126,941 clean reads from SbP, SbPR, and SbW libraries, respectively. Using the filter criteria $|\mathrm{Log_2FC}| \geq 1$ and $P < 0.05$, there were 4,418, 4,898, and 7,378 DEGs in the three comparison groups: SbP vs. SbPR, SbPR vs. SbW, and SbP vs. SbW, respectively. Comparing the three different flowers of *S. baicalensis*, 3,165, 3,457, and 6,026 genes were up-regulated, while 1,253, 1,441, and 1,352 genes were down-regulated in SbP vs. SbPR, SbPR vs. SbW, and SbP vs. SbW, respectively (**Figures 3A–C** and **Supplementary Table 4**). Using GO enrichment analysis, DEGs were enriched in three GO groups: molecular functions, cellular components and biological processes (**Figures 3D–F**). The enrichment analysis of KEGG metabolic pathways of DEGs showed that they were associated with enrichment in metabolic processes, such as

TABLE 1 | Differential accumulation of anthocyanins in the flowers of *Scutellaria baicalensis* 'white petals' (SbW), 'purple-red petals' (SbPR), and 'purple petals' (SbP).

Component name		Content		Fold change	Variable importance in projection (VIP)
SbW_vs_SbP	**Metabolite name**	**SbW**	**SbP**	**SbP/SbW**	
	Cyanidin 3-rutinoside (Keracyanin chloride)	9.00E+00	6.15E+04	6.83E+03	3.01E+00
	Cyanin chloride	8.35E+05	1.99E+07	2.38E+01	1.80E+00
	Cyanidin 3-*O*-galactoside	1.48E+06	3.23E+07	2.19E+01	1.78E+00
	Delphin chloride	4.59E+06	5.41E+07	1.18E+01	1.58E+00
	Peonidin 3-*O*-glucoside chloride	8.73E+04	4.44E+05	5.09E+00	1.29E+00
	Cyanidin 3-*O*-malonylhexoside	8.49E+05	2.60E+07	3.06E+01	1.87E+00
	Cyanidin 3-*O*-glucoside(Kuromanin)	1.78E+06	3.98E+07	2.23E+01	1.79E+00
	Pelargonidin 3-*O*-malonylhexoside	2.84E+03	2.31E+04	8.12E+00	1.46E+00
	Cyanidin *O*-syringic acid	2.14E+05	2.45E+06	1.15E+01	1.58E+00
	Cyanidin *O*-acetylhexoside	1.11E+05	3.83E+06	3.46E+01	1.91E+00
	Centaureidin	9.54E+05	2.99E+06	3.14E+00	1.07E+00
SbW_vs_SbPR	**SbW**	**SbPR**		**SbPR/SbW**	
	Cyanidin 3-rutinoside (Keracyanin chloride)	9.00E+00	9.54E+05	1.06E+05	3.00E+00
	Cyanin chloride	8.35E+05	1.56E+08	1.87E+02	2.02E+00
	Cyanidin 3-*O*-galactoside	1.48E+06	4.55E+07	3.08E+01	1.64E+00
	Delphin chloride	4.59E+06	1.01E+06	2.19E-01	1.09E+00
	Peonidin 3-*O*-glucoside chloride	8.73E+04	6.32E+05	7.24E+00	1.24E+00
	Cyanidin 3-*O*-malonylhexoside	8.49E+05	6.39E+07	7.53E+01	1.84E+00
	Cyanidin 3-*O*-glucoside(Kuromanin)	1.78E+06	5.55E+07	3.12E+01	1.64E+00
	Pelargonidin 3-*O*-malonylhexoside	2.84E+03	5.22E+04	1.84E+01	1.51E+00
	Cyanidin *O*-syringic acid	2.14E+05	1.79E+06	8.36E+00	1.29E+00
	Cyanidin *O*-acetylhexoside	1.11E+05	9.36E+06	8.45E+01	1.86E+00
SbPR_vs_SbP	**SbPR**	**SbP**		**SbP/SbPR**	
	Cyanidin 3-rutinoside (Keracyanin chloride)	9.54E+05	6.15E+04	6.45E-02	1.73E+00
	Cyanin chloride	1.56E+08	1.99E+07	1.27E-01	1.50E+00
	Pelargonin chloride	1.78E+04	3.75E+03	2.10E-01	1.30E+00
	Delphin chloride	1.01E+06	5.41E+07	5.38E+01	2.08E+00
	Centaureidin	3.23E+05	2.99E+06	9.25E+00	1.56E+00

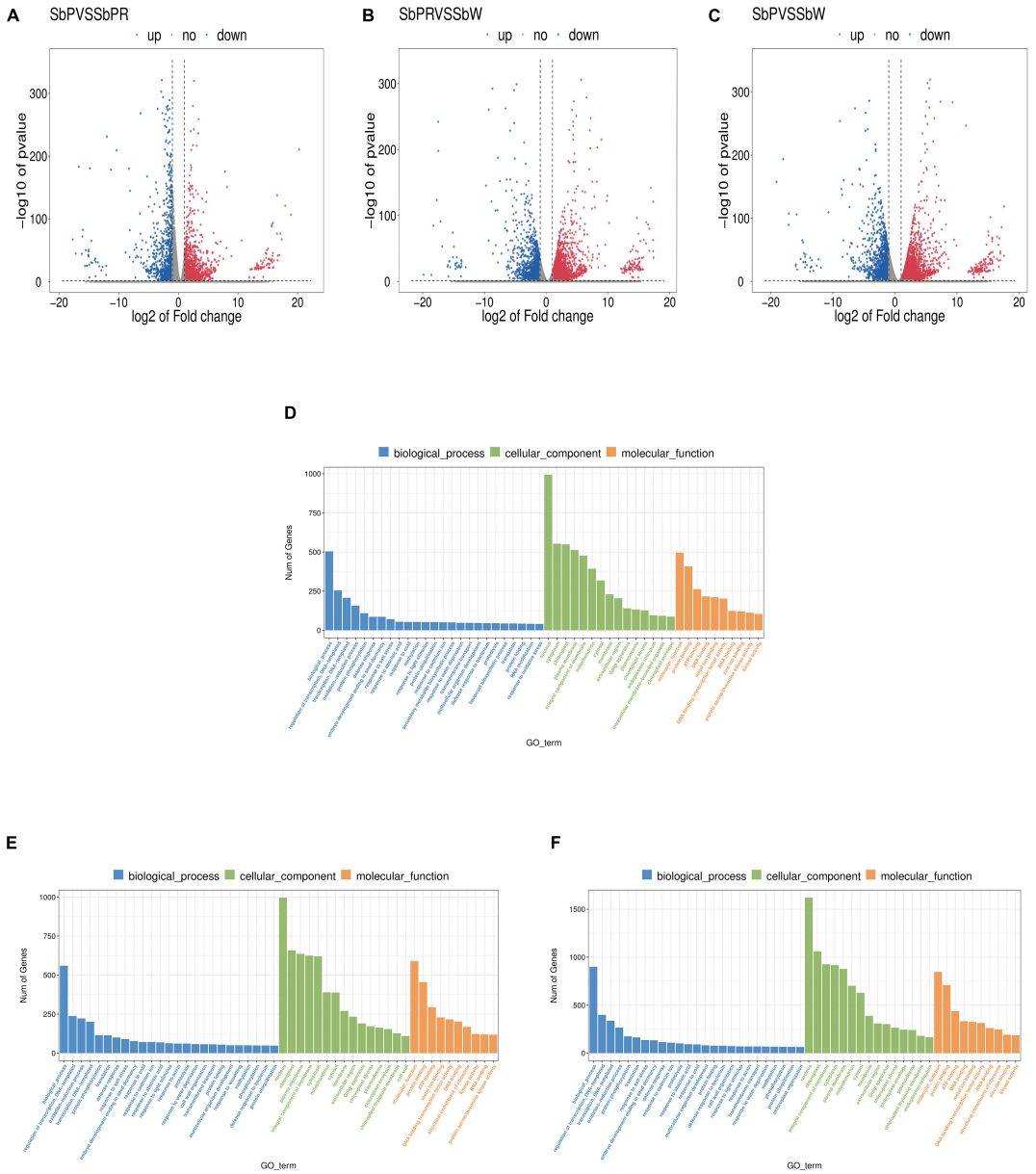


FIGURE 3 | Transcriptome analysis of genes in SbW, SbPR, and SbP. **(A–C)** The volcano plot indicates the differential gene expression levels among SbW, SbPR, and SbP. Blue dots represent down-regulated DEGs; red spots represent up-regulated DEGs; and gray represents non-DEGs. **(D–F)** GO enrichment of the differential genes in SbW, SbPR, and SbP. SbW: white flower at the blooming stage; SbPR, purple-red flower at the blooming stage; SbP, purple flower at the blooming stage.

the flavonoid biosynthesis pathway and phenylalanine synthesis pathway (**Table 2** and **Supplementary Table 5**).

Integrated Transcriptomic and Metabolomic Analyses to Reveal Anthocyanin Synthesis in Three Different Flowers of *Scutellaria baicalensis*

By analyzing the single genes responsible for anthocyanin synthesis, the key enzyme genes of *S. baicalensis* with three different flowers were discovered. Through detailed comparative analysis, it was found that most secondary metabolite pathways were enhanced by the up-regulation of gene expression. In SbW vs. SbP, the expression levels of upstream (CHS, CHI, etc.) and downstream genes (DFR, 3GT, etc.) in the anthocyanin biosynthesis pathway were higher in SbP than in SbW. These genes were increased by 1.01- to 3.9-fold (**Figure 4A**). In SbPR vs. SbP, structural genes CHS, CHI, F3′H, ANS, and DFR increased 1.04 to 4.46-fold, and the Sb3GT gene (an up-regulated and down-regulated DEG) was up-regulated by 2.25-fold and down-regulated by 4.60-fold, respectively. Sb3GT can catalyze the modification of unstable anthocyanins in *S. baicalensis* and

TABLE 2 | Significant enrichment in KEGG pathways among SbW, SbPR, and SbP.

No.	Pathway	DEGs with pathway annotation	All genes with pathway annotation	*P*-value	Pathway ID
SbW_vs_SbP					
1	Anthocyanin biosynthesis	3	23	0.99	ko00942
2	Flavonoid biosynthesis	38	112	0.26	ko00941
3	Phenylpropanoid biosynthesis	73	264	1.78E + 00	ko00940
4	Plant hormone signal transduction	132	572	1.00	ko04075
5	Starch and sucrose metabolism	118	360	0.21	ko00500
SbW_vs_SbPR					
1	Flavonoid biosynthesis	21	112	0.71	ko00941
2	Phenylalanine metabolism	18	92	0.63	ko00360
3	Protein processing in endoplasmic reticulum	59	363	0.98	ko04141
4	Anthocyanin biosynthesis	4	23	0.72	ko00942
SbPR_vs_SbP					
1	Starch and sucrose metabolism	59	360	0.69	ko00500
2	Plant hormone signal transduction	75	572	1.00	ko04075
3	Phenylpropanoid biosynthesis	59	264	0.02	ko00940
4	Flavonoid biosynthesis	29	112	0.01	ko00941
5	Anthocyanin biosynthesis	6	23	0.19	ko00942

transform them into stable anthocyanins (Kobayashi et al., 2001), and the change in its expression may affect the accumulation of anthocyanins. This also supported the increase of metabolite delphin chloride and the decrease of cyanidin 3-rutinoside in SbP (**Figure 4B**). In SbPR vs. SbW, it was found that except for *DFR*, the expression level of other genes in SbPR was lower than that in SbW, but the anthocyanin level was higher in SbPR than in SbW (**Figure 4C**), which suggested that the down-regulated expression of *DFR* affected the formation of anthocyanins. In summary, differentially expressed *DFR* and *Sb3GT* genes can regulate the anthocyanin synthesis pathway in *S. baicalensis*.

Correlation Analysis of Differentially Expressed Genes and Anthocyanins

To analyze the regulatory networks of anthocyanins and genes involved in different flower colors in *S. baicalensis*, three DEGs (*Sb02g31040*, *Sb07g04780*, and *Sb01g72290*) and 15 anthocyanins were selected as the source data for correlation network analysis and heatmap generation. As shown in **Figure 5A**, delphin chloride was highly correlated with *Sb02g31040*(*P_W*), *Sb01g72290*(*P_PR*), *Sb02g31040* (*P_PR*), *Sb07g04780*(*P_PR*), *and Sb07g04780*(*PR_W*), and they were all positively correlated except for *Sb01g72290* (*P_PR*). Cyanidin 3-rutinoside (Keracyanin chloride) was highly correlated with *Sb02g31040*(*P_PR*), *Sb07g04780*(*P_PR*), *Sb02g31040*(*P_W*), *and Sb07g04780*(*PR_W*), and they were all negatively correlated except for *Sb02g31040*(*P_W*) (**Figure 5A**). The transcription factors of DEGs were also identified in the correlation network, including WRKY, WD40, MYB, bHLH, and NAC. Among them, WRKY31, bHLH66, NAC100, and WRKY40 were positively correlated with *Sb01g72290* (*P_PR*), MYB, WD40, WRKY, NAC62, and WRKY53 were negatively correlated with *Sb01g72290* (*P_PR*), NAC79, and NAC2 were negatively correlated with *Sb02g31040* (*P_W*), NAC35 was positively correlated with *Sb02g31040* (*P_W*) (**Figure 5A**). Heatmap analysis showed that there were significant differences between selection DEGs and anthocyanins. Among them, the regulation extent of *Sb02g31040* and *Sb07g04780* to delphin chloride biosynthesis was significantly higher than that of *Sb01g72290*, which was in good agreement with the findings of the correlation network analysis (**Figure 5B**). Furthermore, by analyzing the correlation among anthocyanin content, transcription factors and key genes, we found an interesting phenomenon that TFs positively regulated *Sb01g72290* (*P_PR*) expression resulting in a decrease in anthocyanin accumulation, TFs negatively regulated *Sb01g72290* (*P_PR*) expression resulting in an increase in anthocyanin accumulation, and TFs positively regulated *Sb02g31040* (*P_W*) expression resulting in an increased anthocyanin accumulation (**Figure 6** and **Supplementary Table 6**). In total, *Sb02g31040*, *Sb01g72290* and WRKY, bHLH, and NAC jointly participated in the regulation of the biosynthesis of delphin chloride and cyanidin 3-rutinoside among three different colored flowers of *S. baicalensis*.

Key Genes Responsible for Anthocyanin Biosynthesis in Three Different *Scutellaria baicalensis* Flowers

Based on our transcriptome database analysis, *SbDFR* (Sb02g31040) and two differentially expressed *Sb3GT* genes (Sb07g04780 and Sb01g72290) were identified. BLAST analysis indicated that *SbDFR* has a high similarity (99.87%) with *SvDFR* (GenBank: FJ605512.1). The amino acid sequence analysis showed that the SbDFR protein had two conserved domains in the NADB-Rossmann superfamily, namely the NADPH binding domain and substrate-specific binding domain (**Figure 7A**). Phylogenetic analysis showed that it had the closest genetic relationship with *S. viscidula* (**Figure 7B**). To identify key domains for catalytic activity, I-TASSER program was applied to predict the 3D structure of the DFR protein. The results

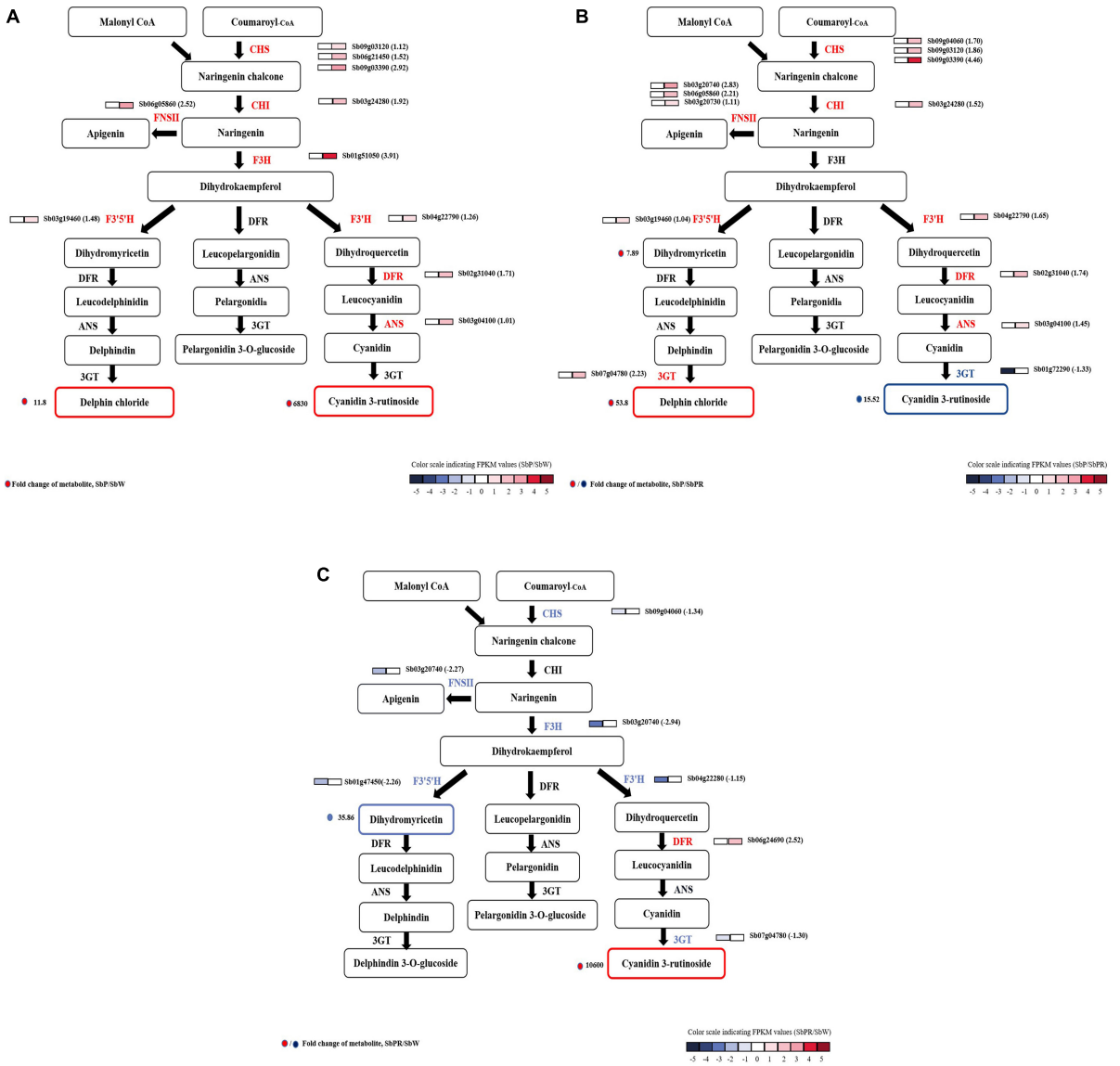


FIGURE 4 | Transcript profiling of genes related to the flavonoid synthesis pathway with different flower colors (SbP, SbPR, and SbW). The core schematic illustration of flavonoid synthesis in *S. baicalensis* flowers was constructed according to KEGG pathway analysis and published literature [**(A)** SbP vs. SbW; **(B)** SbP vs. SbPR; **(C)** SbPR vs. SbW). The annotated flavonoid synthesis genes and the candidate genes for regulating anthocyanins synthesis in *S. baicalensis* flowers are shown in the figure with different colors. The grids with color-scale from dark to light indicate RPKM values; red and blue represent high and low abundances, respectively. Red or dark solid circles represent the log2 fold change in metabolites contents between three groups (SbP vs. SbPR, SbPR vs. SbW, or SbP vs. SbW). ANS, Anthocyanidin synthase; CHI, chalcone isomerise; CHS, chalcone synthase; DFR, dihydroflflavonol 4-reductase; F3′5′H, flavonoid-3′,5′-hydroxylase; F3H, flavanone 3-hydroxylase; FNS, flavone synthase; GT, UDP-glucosyltransferase.

demonstrated that the DFR protein had two pairs of α-helices and six β-turns located at the N-terminal region (aa 1–260). The four predicted α-helices contained amino acids at 24–35, 98–125, 170–188, and 239–251, and six predicted β-turns located at 15–19, 38–45, 88–92, 130–140, 192–197, and 257–260 (**Figure 8**). The six β-corners and two pairs of α-helices form the Rossmann folding region, which is the binding site region of coenzyme $NADP^+$ and DFR substrate specificity. In this study, at site 121, SbW had leucine, which is a hydrophobic amino acid, while SbPR and SbP had serine, which is a hydrophilic amino acid. At site 164, SbPR had lysine, and SbW and SbP had aspartic acid. These two different sites are located in the Rossmann folding structure, and their differences may lead to DFR selectively catalyzing substrates. Based on the 3D model analysis of its key domains, surface catalytically active protein regions are different, and the base mutation in this structural region may affect the change in catalytic substrate.

The Sb3GT belongs to the UDP-glucosyltransferase (UAGT) superfamily. The amino acid sequence analysis showed that Sb3GT (Sb07g04780) had a plant secondary product

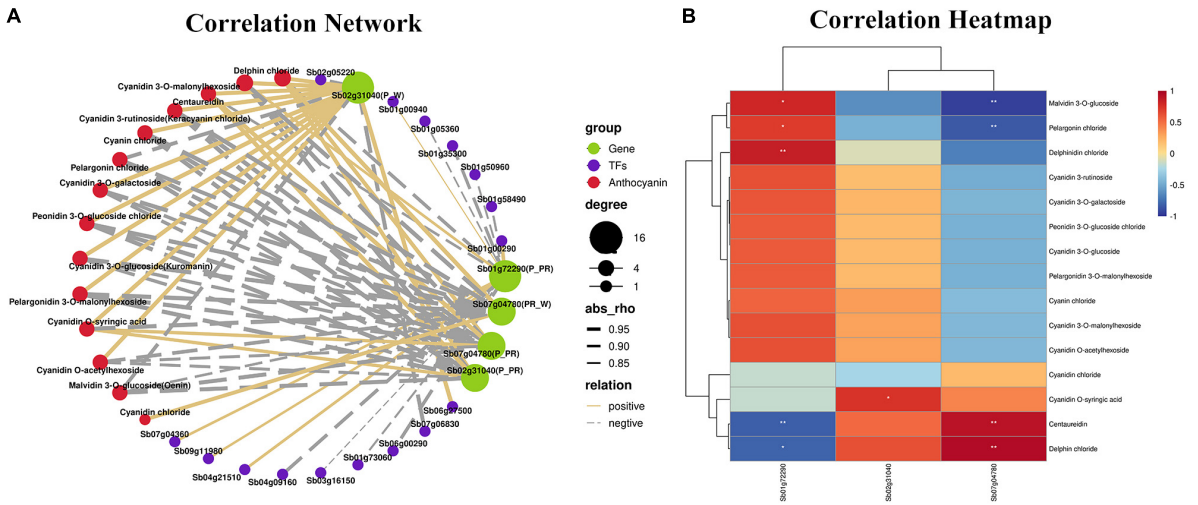


FIGURE 5 | Connection network between metabolite and transcriptome profiles. **(A)** Connection networks among DEGs, TFs, and anthocyanin metabolites. The red, green, and blue colors represents anthocyanin metabolites, DEGs and TFs, respectively. The solid and dotted lines represent positive and negative correlations, respectively. Line width represents the degree of correlation. **(B)** Correlation heatmap of three DEGs in the anthocyanin synthesis pathway and 15 anthocyanins among the three different-colored flowers of *Scutellaria baicalensis*. The heatmap was constructed according to the FPKM value. The red and blue represent positive and negative correlations, respectively.

glycosyltransferase (PSPG) located at the N-terminal region (aa 333–376) (**Figure 9A**). Phylogenetic analysis showed that it had the closest genetic relationship with *S. miltiorrhiza* (**Figure 9B**). The sequence of *Sb3GT* (Sb01g72290) was not found in the NCBI databases. It is speculated that this gene is a novel *Sb3GT* gene in *S. baicalensis*.

The correlation between the expression level of these genes and total anthocyanin content was further calculated. The results showed that two DEGs positively regulated the anthocyanin synthesis, and one DEG negatively regulated anthocyanin synthesis. The expression levels of two DEGs, namely *SbDFR* (Sb02g31040) and *Sb3GT* (Sb07g04780), indicated a significant positive correlation with total anthocyanin content in the samples, suggesting that these two DEGs may have an essential role in anthocyanin accumulation (**Supplementary Table 7**).

Transcription Factors Associated With the Accumulation of Anthocyanin

Transcription factors are involved in the biosynthesis of anthocyanins in plants by affecting target gene expression. By comparing SbP vs. SbW, SbP vs. SbPR, and SbPR vs. SbW, 127, 79, and 73 TFs-related DEGs, respectively, were identified and were mainly annotated as bHLH, MYB, WD40, WRKY, NAC, and MADS-box. These transcription factors may separate or form a complex to regulate the biosynthesis of anthocyanin metabolites in the three different flowers of *S. baicalensis* (**Supplementary Table 8**). Furthermore, by calculating the correlation between the expression levels of these TFs and total anthocyanins content, 81 TFs with remarkable correlation ($|Cor| \geq 0.5$, P value < 0.05) were involved in the accumulation of anthocyanins were identified, including 13 negative regulator and 68 positive regulators. These negative regulators include 3 WRKY, 1 NAC, 5 MYB, and 4 MADS box, which likely act as a repressor in anthocyanin synthesis. The 68 positive regulators, including 3 WRKY (such as WRKY53), 11 WD40, 4 NAC (such as NAC35), 14 MYB, 5 MADS box, 1 bZIP, 11 bHLH, and 19 others, may act as promotors of anthocyanin accumulation (**Supplementary Table 9**).

Validation of the Transcriptomic Data Through Quantitative Real-Time Polymerase Chain Reaction

To evaluate the reliability of transcriptome information, 12 DEGs (3 transcription factor genes and nine flavonoid biosynthetic pathway genes) were chosen to validate the RNA-seq data. The qRT-PCR results indicated that the expression profiles of 12 selected DEGs were highly correlated with those obtained from the transcriptome data (**Figure 10**).

DISCUSSION

Identification of Anthocyanin From the Flowers of *Scutellaria baicalensis*

Anthocyanins, as important secondary metabolites, widely exist in flowers, fruit, leaves, and seed coats and participate in many physiological and biochemical reactions of plants (Liu et al., 2001; The Angiosperm Phylogeny Group, 2009; Dong et al., 2011; Brunetti et al., 2013; Mierziak et al., 2014; Grimes et al., 2018). In the present study, to explore anthocyanin synthesis pathway in white, purple, and purple-red flowers of *S. baicalensis*, transcriptomic and metabolomic analyses were conducted. Based on UPLC-MS, 168 metabolites were identified, including 15 anthocyanins. The analysis of anthocyanin differential

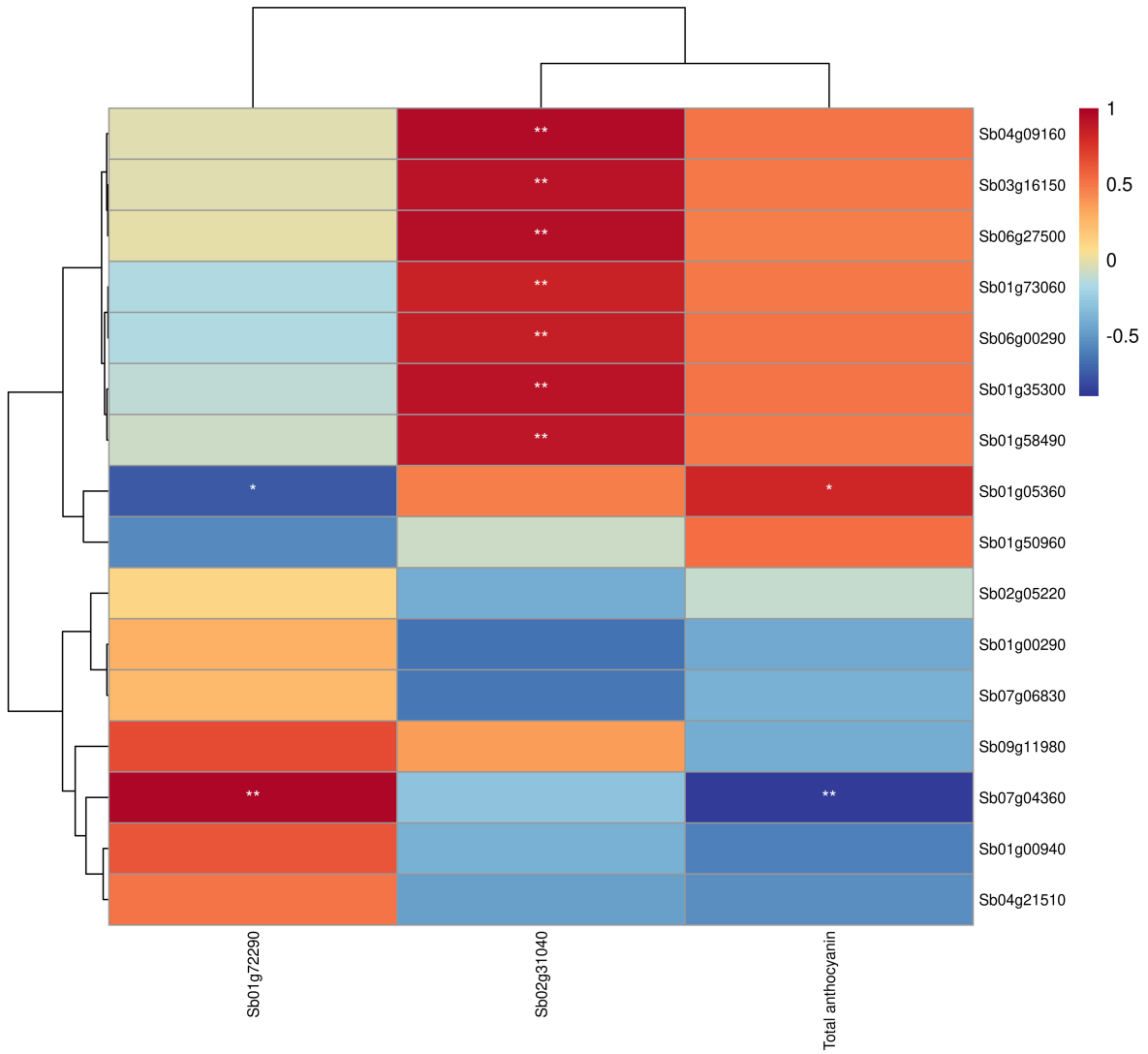


FIGURE 6 | Correlation heatmap of TFs with total anthocyanins and key genes. The color scale on the right represents the degree of correlation among TFs, genes and anthocyanins. The red represents high correlation degree and blue represents low correlation degree. One asterisk represents correlation and two asterisks represent significant correlation.

metabolites in SbP, SbPR, and SbW showed that the content of cyanidin 3-rutin (chlorokeratin), reported as the major anthocyanins in the peels of "Brown Turkey", "Bursa", and "Black Mission" figs (Solomon et al., 2006; Ercisli et al., 2012), was 105,947- and 6,828-fold higher in SbPR and SbP than in SbW, respectively (**Table 1**). The cyanidin 3-rutinoside (keracyanin chloride) content was 15.52-fold higher in SbPR vs. SbP, while the content of delphin chloride in SbPR was 53.78-fold lower than in SbP. This result suggests that cyanidin 3-rutinoside and delphin chloride are the key anthocyanins conferring pigment accumulation in flowers of *S. baicalensis*. Interestingly, this study found that there were six common anthocyanin pigments (cyanidin, delphinidin, pelargonidin, petunidin, peonidin, and malvidin) present in *S. baicalensis* flowers. The decisive factors of flower color may be related to the number of pigment molecules, metallic ions and/or various molecular conformations of anthocyanins. The pigment molecules accumulated to different degrees in *S. baicalensis* flowers, which induced the biosynthesis of anthocyanins with different colors in the metabolic pathway.

Genes Involved in Anthocyanin Biosynthesis

Through transcriptome sequencing analysis and the identification of single genes involved in the biosynthesis of anthocyanins, CHI, CHS, FNSII, F3′H, F3H, F3′5′H, 3GT, and

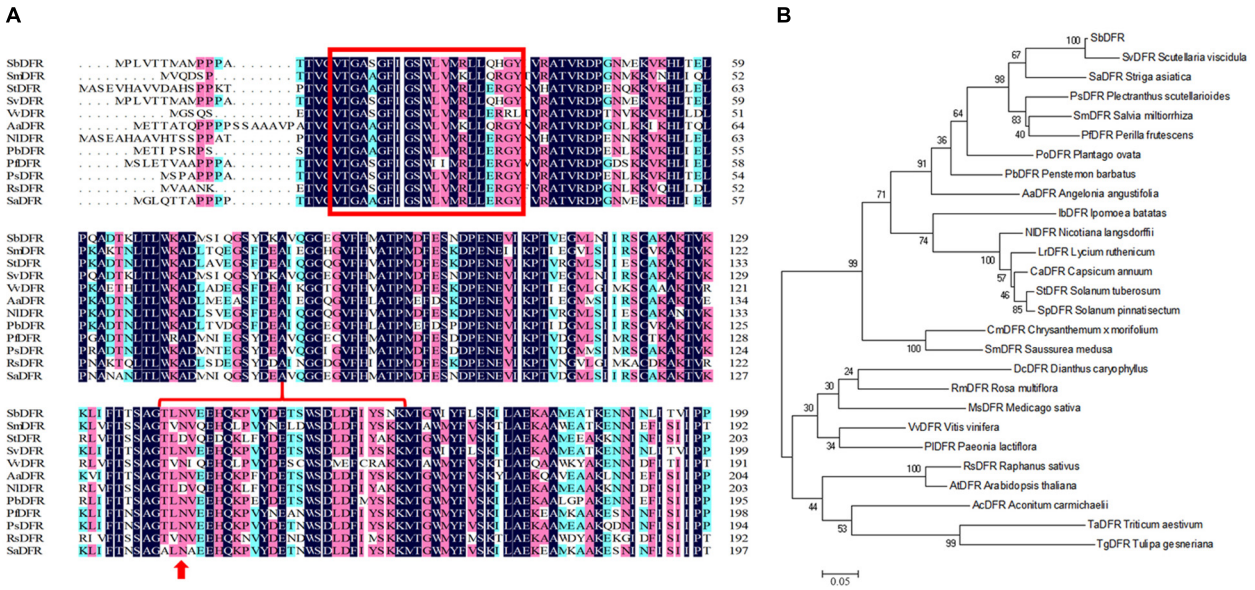


FIGURE 7 | Phylogenetic and multiple alignment analyses of the deduced amino acid sequence of the SbDFR protein from *Scutellaria baicalensis* and other species. **(A)** Multiple alignment of the proteins from selected species. The NADPH binding domain and substrate specific binding domain are marked by the red box and red brace, respectively, and the site marked by the red arrow is the amino acid residue influencing enzyme substrate specificity. **(B)** Phylogenetic analysis of SbDFR proteins from *S. baicalensis* and other selected species.

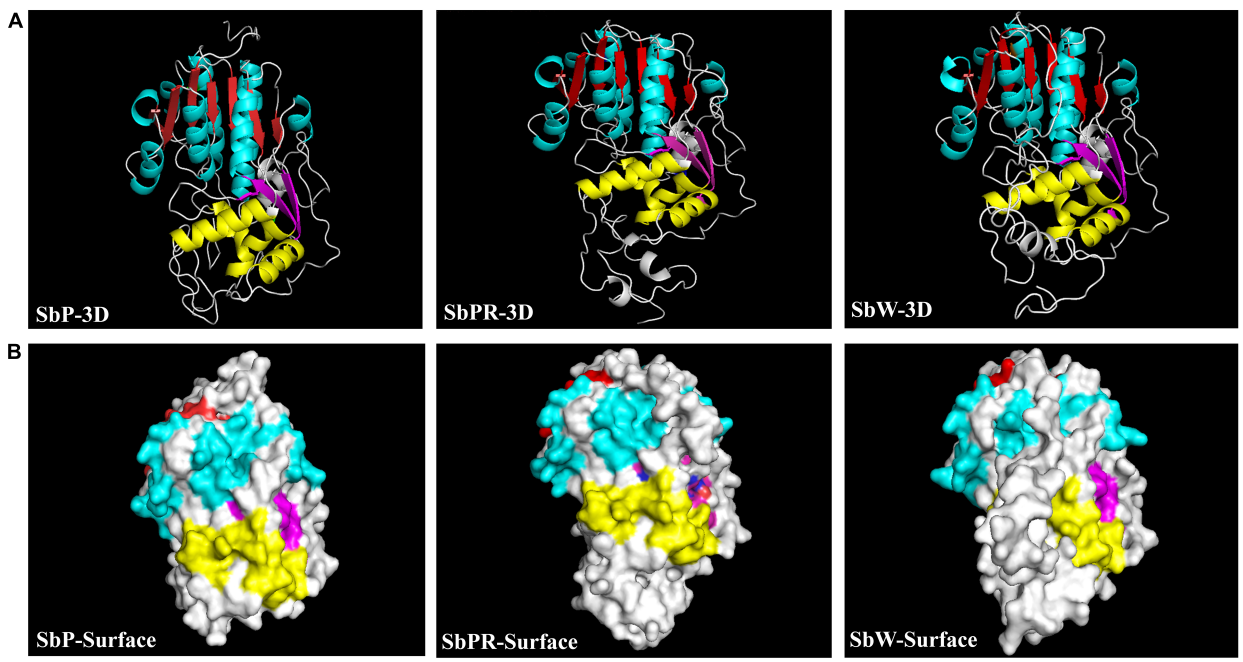


FIGURE 8 | The predicted 3D protein structure of DFR in *Scutellaria baicalensis* flowers. **(A)** The 3D structural model of the full-length DFR in SbP, SbPR, and SbW. **(B)** Predicted surface localization corresponding to three DFR secondary structures were revealed by UCSF Chimera program. The same secondary structure in the same area is marked with the same color. The red and yellow marks represent the same secondary structure of α-helice in different functional areas, and the blue and purple marks represent the same secondary structure of β-turn in different functional areas.

DFR were expressed to different degrees in three different flower colors of *S. baicalensis*. In SbPR vs. SbP, the structural gene *Sb3GT* in the anthocyanin synthesis pathway was differentially expressed, while in SbPR vs. SbW, other significant structural genes except the *SbDFR* gene were down-regulated. Studies have found that the expression of enzymes-encoding genes is involved

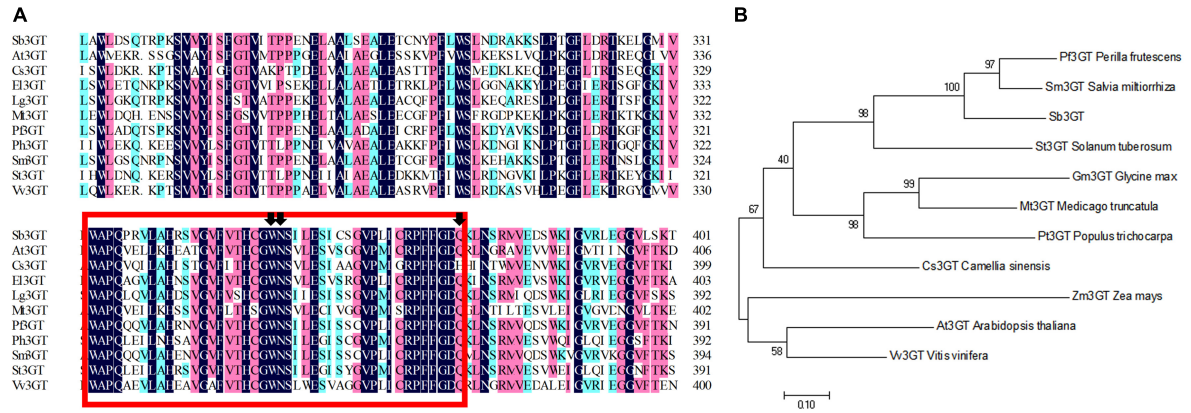


FIGURE 9 | Phylogenetic and multiple alignment analyses of the deduced amino acid sequence of Sb3GT protein from *Scutellaria baicalensis* and other species. **(A)** Multiple alignment of the proteins from selected species. A N-terminal-conserved plant secondary product glycosyltransferases (PSPG) domain is framed in a red box. The black arrow indicates the significant amino acid sites of 22, 23, and 44 in the PSPG domain. **(B)** Phylogenetic analysis of the Sb3GT proteins from *S. baicalensis* and other selected species.

in the regulation of anthocyanin synthesis (Saito et al., 2013). For example, high expression of *ANS*, *DFR* and anthocyanidin 3-*O*-glucosyltransferase (*UFGT*) can affect color changes in fruit (Han et al., 2012). The down-regulated expression levels of *DFR* and *ANS* genes inhibit the synthesis of anthocyanins, thus resulting in the formation of white flowers. According to the data of metabolomics and transcriptomics, the low expression of *SbDFR* reduced the anthocyanin levels in *S. baicalensis* flowers, and the differential expression of *Sb3GT* led to the formation of purple flower and purple-red flower of *S. baicalensis*.

Function of *SbDFR* and *Sb3GT* Genes

Anthocyanin synthesis in plants is affected by a series of key enzyme genes and related transcription factors (Suzuki et al., 2016). It has been reported that the down-regulation of *DFR* and *ANS* genes in the anthocyanin synthesis pathway leads to pigmentation loss (Bogs et al., 2007; Clark and Verwoerd, 2011). The over-expression of different *DFR* genes in tobacco flowers promotes the biosynthesis of anthocyanins and increases the deposition of red pigments (Luo et al., 2016). Our study demonstrated that the white-flowered characteristic was attributed to the down-regulated expression of *DFR*. Anthocyanins are highly unstable and easily degraded, so glycosylation is very important to stabilize anthocyanins. 3GT belongs to the flavonoid glycosyltransferase family and has a highly conserved PSPG box. It can catalyze the modification of unstable anthocyanins in plants and transform them into stable anthocyanins (Wu et al., 2017). The 3GT gene of plum blossom has higher activity in red flowers than in white flowers, and the enzyme activity increases with the appearance of red coloring (Niu et al., 2010). The accumulation of anthocyanins in *Myrica rubra* fruit is related to the coordinated expression of many biosynthetic genes (DFR, ANS, F3′H, F3H, and 3GT) and is regulated by MrMYB1 (Wu et al., 2017). Our study identified two *Sb3GT* genes (*Sb07g04780 and Sb01g72290*), and the differential expression of *Sb3GT* led to the formation of purple and purple-red flowers in *S. baicalensis*. However, we speculated that the *Sb3GT* (Sb01g72290) gene is a new 3GT gene in *S. baicalensis*, which functions in purple and purple-red flowers in *S. baicalensis* and needs further study.

Transcription Factors

Studies have shown that TFs play an essential role in the modulation of anthocyanin synthesis pathways. For example, FtMYB1 and FtMYB2 in tartary buckwheat and VvMYBA2 in grapes can promote the expression of *DFR* (Bai et al., 2014; Niu et al., 2016). In apples, the *CHS* gene is positively correlated with the expression of MYB4 and MYB5 (Clark and Verwoerd, 2011). Abnormal expression of bHLH3 disrupts the balance of the flavonoid regulatory network required for fruit development, leading to differences in pigment composition in mulberry fruit (Lloyd et al., 2017). MYB TFs (e.g., PAP1/PAP2, TT2, MYB-75, -90, -113, and 114), bHLH TFs (e.g., EGL3, GL3, and TT8), and WD40 repeat protein (TTG1) in *Arabidopsis thaliana* form an MBW complex to participate in regulating the expression of *UFGT*, *ANS*, *DFR*, and other downstream genes, thus affecting the biosynthesis of anthocyanins (Xu et al., 2014). Apart from MYB, bHLH, and WD40, transcription factors of zinc fingers, MADs, and WRKY proteins are also responsible for regulating the biosynthesis of anthocyanins (Terrier et al., 2009; Li H. et al., 2020). Besides, *IbMADS10* modulates the biosynthesis of anthocyanins to enhance anthocyanin pigment accumulation in sweet potato (Lalusin et al., 2011). NAC TFs (e.g., CUC2, ATAF-1, -2, and NAM) have shown to regulates biosynthesis of anthocyanins in blood-fleshed peaches (Zhou et al., 2015). It has been reported that MYB114 and MYB75 play a key role in the modulation of anthocyanin synthesis pathways. Suppression of PyMYB114 could inhibit anthocyanin biosynthesis in red-skinned pears (Yao et al., 2017). Co-transformation of Pp12ERF96 with PpMYB114 and PpbHLH3 in tobacco leaves led to enhanced anthocyanin accumulation (Ni et al., 2019). MYB75 defined as PRODUCTION OF

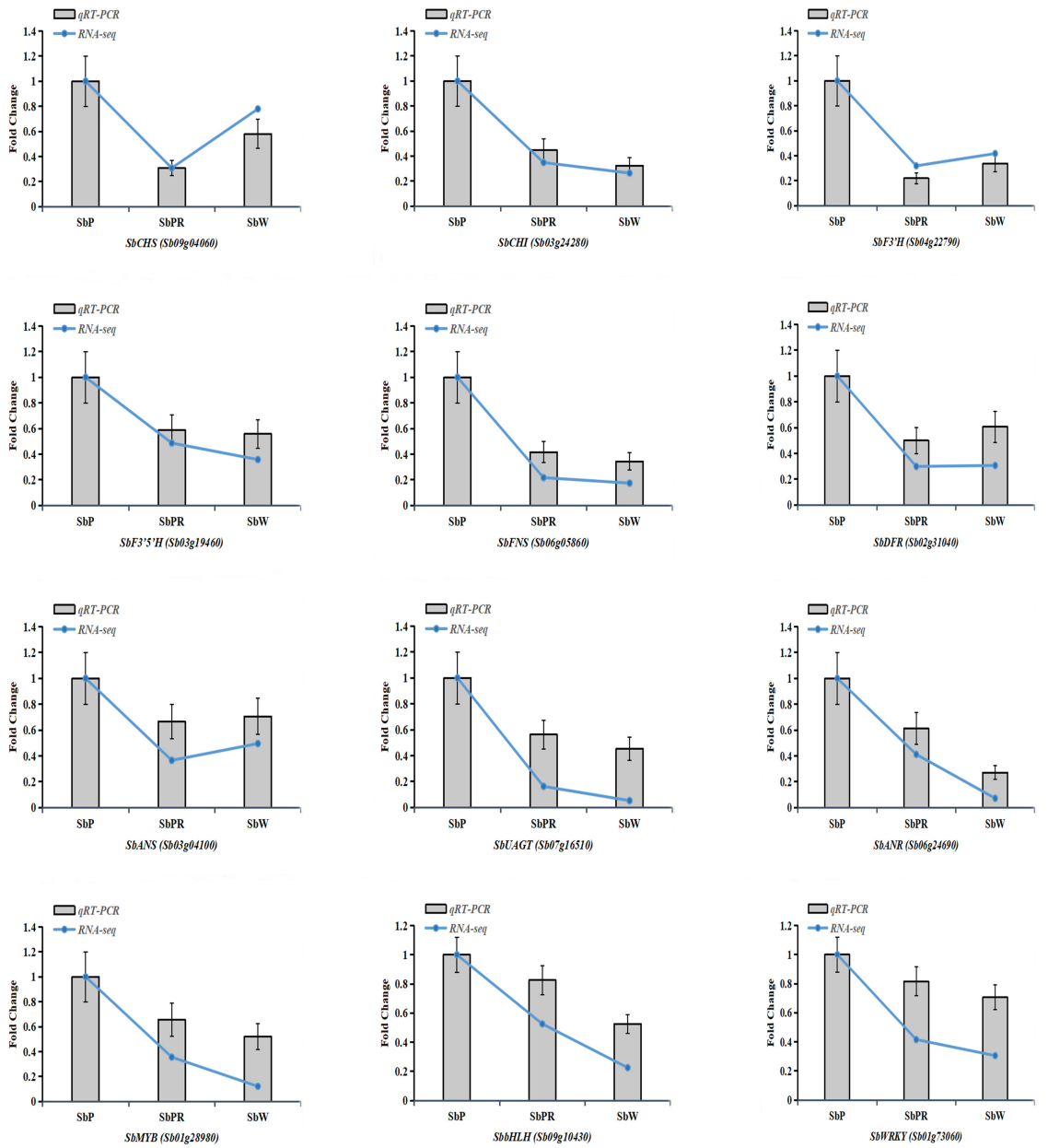


FIGURE 10 | Verification of representative genes related to anthocyanin biosynthesis by qRT-PCR. Relative expression levels of candidate genes were calculated using the *actin* gene as a standard. Error bars represent the standard deviation of three replicates. Curve graphs represent gene expression from the RNAseq data based on FPKM. Bar charts represent the gene expression levels calculated from qRT-PCR data based on ΔCt method.

ANTHOCYANIN PIGMENT 1 (PAP1), plays a key role in anthocyanin accumulation (Antonio et al., 2008). In this study, after analysis and comparison of the *Scutellaria baicalensis* petal transcriptome sequencing database, no related genes with high homology of MYB114 and MYB75 were found, but we found that myb family transcription factor APL isoform X3 (**Supplementary Table 6**) was significantly positively correlated with the accumulation of anthocyanins. It is speculated that myb family transcription factor APL isoform X3 can play a key role in the synthesis of anthocyanin of *Scutellaria baicalensis*, which its regulatory role may similar to the reported TFs of MYB114 and myb75, but the specific function needs to be further verified by more experiments.

In our study, 127, 79, and 73 TFs-related DEGs were identified in SbP vs. SbW, SbP vs. SbPR, and SbPR vs. SbW, respectively, including bHLH, MYB, WRKY, WD40, NACs, and MADS-box (**Supplementary Table 8**). We speculate that these TFs-related DEGs (such as WRKY, bHLH, and NAC) can serve as key players of anthocyanin synthesis in *S. baicalensis* flowers. Correlation analysis also revealed that the TFs WRKY, bHLH,

and NAC were closely related to anthocyanin accumulation, gene expression and anthocyanin synthesis in different degree of *S. baicalensis* flowers, but their patterns of regulating gene expression and metabolite biosynthesis are unclear. Therefore, further studies are needed to verify whether these candidate TFs can form a ternary protein complex or alone directly affects anthocyanin synthesis in *S. baicalensis* flowers.

CONCLUSION

To elucidate the underlying mechanism of color differentiation in metabolic and transcriptional differences among three kinds of *S. baicalensis* flowers, metabolic, transcriptional, and integration analyses were performed, and the DEGs responsible for the biosynthesis and regulation of anthocyanins were identified. Our results showed that cyanidin 3-rutin and delphinium chloride metabolites were considered to be the key anthocyanins for pigment accumulation in *S. baicalensis* flowers, and the expression levels of *SbDFR* and *Sb3GT* genes were strongly correlated with the anthocyanin biosynthesis process in *S. baicalensis* flowers. In addition, transcription factors such as WRKY, bHLH, and NAC, especially for NAC35, which were also positively involved in anthocyanin biosynthesis in *S. baicalensis* flowers. Our results will provide a solid foundation for understanding the roles of DEGs and TFs in regulating anthocyanin biosynthesis in *S. baicalensis* flowers and also guide future directional breeding of new *S. baicalensis* germplasm for enriched agricultural landscape design.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA811392; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA811392.

AUTHOR CONTRIBUTIONS

DW, JW, and YN conceived and designed the experiments. JW and YW collected the plant samples. JW, DY, and YW performed the experiments. DY conducted bioinformatics analysis. DW and JW wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This research was funded by the Natural Science Foundation of Shanxi Province of China (Grant No. 20210302123376), China Agriculture Research System of MOF and MARA (Grant No. CARS-21) and the earmarked fund for Modern Agro-industry Technology Research System (Grant No. 2021-11).

ACKNOWLEDGMENTS

We would like to thank Wuhan Metware Biotechnology Co., Ltd. and Hangzhou Lianchuan Biotechnology Co., Ltd., for assisting in bioinformatics analyzing and sequencing. We would also like to express their gratitude to LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.884957/full#supplementary-material

Supplementary Figure 1 | The heatmap analysis of all differentially accumulated metabolites among the flowers of SbW, SbPR, and SbP. The color scale on the right represents the expression degree of differentially accumulated flavonoid metabolites in the material. Red represents high accumulation level and blue represents low accumulation level.

Supplementary Figure 2 | The heatmap analysis of all differentially accumulated anthocyanins among the flowers of SbW, SbPR, and SbP according to the relative content in samples. The color scale on the right represents the expression degree of differentially accumulated anthocyanin metabolites in the material. Red represents high accumulation level and blue represents low accumulation level.

Supplementary Figure 3 | Number of differentially accumulated anthocyanins among SbPR vs SbP, SbW vs SbP, and SbW vs SbPR.

REFERENCES

Antonio, G., Zhao, M. Z., John, M. L., and Alan, M. L. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J.* 53, 814–827.

Bai, S. L., Tao, R. Y., Yin, L., Ni, J. B., Yang, Q. S., Yan, X. H., et al. (2019). Two B-box proteins, PpBBX18 and PpBBX21, antagonistically regulate anthocyanin biosynthesis via competitive association with pyrus pyrifolia elongated hypocotyl 5 in the peel of pear fruit. *Plant J.* 100, 1208–1223. doi: 10.1111/tpj.14510

Bai, Y. C., Li, C. L., Zhang, J. W., Li, S. J., Luo, X. P., Yao, H. P., et al. (2014). Characterization of two tartary buckwheat R2R3-MYB transcription factors and their regulation of proanthocyanidin biosynthesis. *Physiol. Plant.* 152, 431–440. doi: 10.1111/ppl.12199

Baumann, S., Fas, S. C., Giaisi, M., Müller, W. W., Merling, A., Gülow, K., et al. (2008). Wogonin preferentially kills malignant lymphocytes and suppresses T-cell tumor growth by inducing PLCg1- and $Ca2^{+}$-dependent apoptosis. *Blood* 111, 2354–2363. doi: 10.1182/blood-2007-06-096198

Bogs, J., Jaffe, F. W., Takos, A. M., Walker, A. R., and Robinson, S. P. (2007). The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. *Plant Physiol.* 143, 1347–1361. doi: 10.1104/pp.106.093203

Brunetti, C., Di Ferdinando, M., Fini, A., Pollastri, S., and Tattini, M. (2013). Flavonoids as antioxidants and developmental regulators: relative significance

in plants and humans. *Int. J. Mol. Sci.* 14, 3540–3555. doi: 10.3390/ijms14023540

Chen, K. L., Du, L. J., Liu, H. L., and Liu, Y. L. (2019). A novel R2R3-MYB from grape hyacinth, MaMybA, which is different from MaAN2, confers intense and magenta anthocyanin pigmentation in tobacco. *BMC Plant Biol.* 19:390. doi: 10.1186/s12870-019-1999-0

Chen, Q., Yu, H., Tang, H., and Wang, X. (2012). Identification and expression analysis of genes involved in anthocyanin and proanthocyanidin biosynthesis in the fruit of blackberry. *Sci. Horticult.* 141, 61–68.

Chen, W., Gong, L., Guo, Z. L., Wang, W. S., Zhang, H. Y., Liu, X. Q., et al. (2013). A novel integrated method for large-scale detection, identifification, and quantifification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Cheng, H., Li, L. L., Cheng, S. Y., Cao, F. L., Xu, F., Yuan, H. H., et al. (2013). Molecular cloning and characterization of three genes encoding dihydroflavonol-4-reductase from Ginkgo biloba in anthocyanin biosynthetic pathway. *PLoS One* 8:e72017. doi: 10.1371/journal.pone.0072017

Clark, S. T., and Verwoerd, W. S. (2011). A systems approach to identifying correlated gene targets for the loss of colour pigmentation in plants. *BMC Bioinformat.* 12:343. doi: 10.1186/1471-2105-12-343

Dong, T. T., Han, R. P., Yu, J. W., Zhu, M. K., Zhang, Y., Gong, Y., et al. (2019). Anthocyanins accumulation and molecular analysis of correlated genes by metabolome and transcriptome in green and purple asparaguses (*Asparagus officinalis*, L.). *Food Chem.* 271, 18–28.

Dong, Y., Ji, G., Cao, A. L., Shi, J. R., Shi, H. L., Xie, J. Q., et al. (2011). Effects of sinensetin on proliferation and apoptosis of human gastric cancer AGS cells. *Chin. J. Tradit. Chin. Med.* 36, 790–794.

Dooner, H. K., Robbins, T. P., and Jorgensen, R. A. (1991). Genetic and developmental control of anthocyanin biosynthesis. *Annu. Rev. Genet* 25, 173–199. doi: 10.1146/annurev.ge.25.120191.001133

Duan, H. R., Wang, L. R., Cui, G. X., Zhou, X. H., Duan, X. R., and Yang, H. S. (2020). Identification of the regulatory networks and hub genes controlling alfalfa floral pigmentation variation using RNA-sequencing analysis. *BMC Plant Biol.* 20:110. doi: 10.1186/s12870-020-2322-9

Ercisli, S., Tosun, M., Karlidag, H., Dzubur, A., Hadziabulic, S., and Aliman, Y. (2012). Color and antioxidant characteristics of some fresh fig (*Ficus carica* L.) genotypes from Northeastern Turkey. *Plant Foods Hum. Nutr.* 67, 271–276. doi: 10.1007/s11130-012-0292-2

Fraga, C. G., Clowers, B. H., Moore, R. J., and Zink, E. M. (2010). Signature discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and chemometrics. *Anal. Chem.* 82, 4165–4173. doi: 10.1021/ac1003568

Gao, J. Y., Morgan, W. A., Sanchez-Medina, A., and Corcoran, O. (2011). The ethanol extract of Scutellaria baicalensis and the active compounds induce cell cycle arrest and apoptosis including upregulation of p53 and Bax in human lung cancer cells. *Toxicol. Appl. Pharmacol.* 254, 221–228. doi: 10.1016/j.taap.2011.03.016

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

Grimes, K. L., Stuart, C. M., McCarthy, J. J., Kaur, B., Cantu, E. J., and Forester, S. C. (2018). Enhancing the cancer cell growth inhibitory effects of table grape anthocyanins. *J. Food Sci.* 83, 2369–2374. 14294 doi: 10.1111/1750-3841

Han, Y., Vimolmangkang, S., Soria-Guerra, R. E., and Korban, S. S. (2012). Introduction of apple ANR genes into tobacco inhibits expression of both CHI and DFR genes in flowers, leading to loss of anthocyanin. *J. Exp. Bot.* 63, 2437–2447. doi: 10.1093/jxb/err415

He, J., and Giusti, M. M. (2010). Anthocyanins: natural colorants with health-promoting properties. *Annu. Rev. Food Sci. Technol.* 1, 163–187. doi: 10.1146/annurev.food.080708.100754

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Jiang, T., Zhang, M. D., Wen, C. X., Xie, X. L., Tian, W., Wen, S. Q., et al. (2020). Integrated metabolomic and transcriptomic analysis of the anthocyanin regulatory networks in Salvia miltiorrhiza Bge. flowers. *BMC Plant Biol.* 20:349. doi: 10.1186/s12870-020-02553-7

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008). KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* 36, 480–484. doi: 10.1093/nar/gkm882

Kobayashi, S., Ishimaru, M., Ding, C. K., Yakushiji, H., and Goto, N. (2001). Comparison of UDP-glucose: flavonoid 3-O-glucosyltransferase (UFGT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin. *Plant Sci.* 160, 543–550. doi: 10.1016/s0168-9452(00)00425-8

Lalusin, A. G., Ocampo, E. T. M., and Fujimura, T. (2011). Arabidopsis thaliana plants over-expressing the IbMADS10 gene from sweetpotato accumulates high level of anthocyanin. *Philipp. Agric. Sci.* 36, 30–36.

Li, D., Chen, G., Ma, B., Zhong, C. G., and He, N. J. (2020). Metabolic profiling and transcriptome analysis of mulberry leaves provide insights into flavonoid biosynthesis. *J. Agric. Food Chem.* 68, 1494–1504. doi: 10.1021/acs.jafc.9b06931

Li, H., Yang, Z., Zeng, Q. W., Wang, S. B., Luo, Y. W., Huang, Y., et al. (2020). Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits. *Hortic. Res.* 7, 83–101. doi: 10.1038/s41438-020-0302-8

Li, P., Ma, F., and Cheng, L. (2013). Primary and secondary metabolism in the sun-exposed peel and the shaded peel of apple fruit. *Physiol. Plant.* 148, 9–24. doi: 10.1111/j.1399-3054.2012.01692.x

Li, S. Z. (2012). *Compendium of Materia Medica (Bencao Gangmu)*. China: Huaxia Press, 543–546.

Liu, X. M., Xiao, G. S., and Chen, W. D. (2001). Advances in research and development of mulberry. *Chin. Tradit. Herbal Drugs* 32, 569–571.

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., et al. (2017). Advances in the MYB–bHLH–WD repeat (MBW) pigment regulatory model: addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant Cell Physiol.* 58, 1431–1441. doi: 10.1093/pcp/pcx075

Lou, Q., Liu, Y., Qi, Y., Jiao, S., Tian, F., Jiang, L., et al. (2014). Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J. Exp. Bot.* 65, 3157–3164. doi: 10.1093/jxb/eru168

Luo, P., Ning, G. G., Wang, Z., Shen, Y. X., Jin, H. N., Li, P. H., et al. (2016). Disequilibrium of flavonol synthase and dihydroflavonol-4- reductase expression associated tightly to white vs. red color flower formation in plants. *Front. Plant Sci.* 6:1257. doi: 10.3389/fpls.2015.01257

Meng, J., Wang, B., He, G., Wang, Y., Tang, X. F., Wang, S. M., et al. (2019). Metabolomics integrated with transcriptomics reveals redirection of the phenylpropanoids metabolic flux in Ginkgo biloba. *J. Agric. Food Chem.* 67, 3284–3291. doi: 10.1021/acs.jafc.8b06355

Mierziak, J., Kostyn, K., and Kulma, A. (2014). Flavonoids as important molecules of plant interactions with the environment. *Molecules* 19, 16240–16265. doi: 10.3390/molecules191016240

Nayak, M. K., Agrawal, A. S., Bose, S., Naskar, S., Bhowmick, R., Chakrabarti, S., et al. (2014). Antiviral activity of baicalin against influenza virus H1N1-pdm09 is due to modulation of NS1-mediated cellular innate immune responses. *J. Antimicrob. Chemother.* 69, 1298–1310. doi: 10.1093/jac/dkt534

Ni, J., Bai, S., Zhao, Y., Qian, M., Tao, R., Yin, L., et al. (2019). Ethylene response factors Pp4ERF24 and Pp12ERF96 regulate blue light-induced anthocyanin biosynthesis in 'Red Zaosu'pear fruits by interacting with MYB114. *Plant Mol. Biol.* 99, 67–78. doi: 10.1007/s11103-018-0802-1

Niu, S. S., Xu, C. J., Zhang, W. S., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (myrica rubra) fruit by a R2R3 MYB transcription factor. *Planta* 231, 887–899. doi: 10.1007/s00425-009-1095-z

Niu, T. Q., Gao, Z. D., Zhang, P. F., Zhang, X. J., Gao, M. Y., Ji, W., et al. (2016). MYBA2 gene involved in anthocyanin and flavonol biosynthesis pathways in grapevine. *Genet. Mol. Res.* 15, 1–12. doi: 10.4238/gmr15048922

Peng, J. P., Dong, X. J., Xue, C., Liu, Z. M., and Cao, F. X. (2021). Exploring the molecular mechanism of blue flower color formation in Hydrangea macrophylla cv."Forever Summer". *Front. Plant Sci.* 12:585665. eCollection 2021 doi: 10.3389/fpls.2021.585665

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

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., et al. (2013). The flavonoid biosynthetic pathway in Arabidopsis:

structural and genetic diversity. *Plant Physiol. Biochem.* 72, 21–34. doi: 10.1016/j.plaphy.2013.02.001

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C (T) method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Shang, X. F., He, X. R., He, X. Y., Li, M. X., Zhang, R. X., Fan, P. C., et al. (2010). The genus Scutellaria an ethnopharmacological and phytochemical review. *J. Ethnopharmacol* 128, 279–313. doi: 10.1016/j.jep.2010.01.006

Silva, F. M., Bailón, M. T. E., Alonso, J. J. P., Rivas-Gonzalo, J. C., and Santos-Buelga, C. (2007). Anthocyanin pigments in strawberry. *LWT-Food Sci. Technol.* 40, 374–382. doi: 10.1016/j.lwt.2005.09.018

Solomon, A., Golubowicz, S., Yablowicz, Z., Grossman, S., Bergman, M., Gottlieb, H. E., et al. (2006). Antioxidant activities and anthocyanin content of fresh fruits of common fig (*Ficus carica* L.). *J. Agric. Food Chem.* 54, 7717–7723. doi: 10.1021/jf060497h

Suzuki, K., Suzuki, T., Nakatsuka, T., Dohra, H., Yamagishi, M., Matsuyama, K., et al. (2016). RNA-seq-based evaluation of bicolor tepal pigmentation in asiatic hybrid lilies (lilium spp.). *BMC Genomics* 17:611. doi: 10.1186/s12864-016-2995-5

Tanaka, Y., Brugliera, F., and Chandler, S. (2009). Recent progress of flower colour modification by biotechnology. *Int. J. Mol. Sci* 10, 5350–5369. doi: 10.3390/ijms10125350

Terrier, N., Torregrosa, L., Ageorges, A., Vialet, S., Verries, C., Cheynier, V., et al. (2009). Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. *Plant Physiol.* 149, 1028–1041. doi: 10.1104/pp.108.131862

The Angiosperm Phylogeny Group (2009). An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Botan. J. Linnean Soc.* 161, 105–121. doi: 10.1016/j.jep.2015.05.035

Wang, H. X., Fan, W. J., Li, H., Yang, J., Huang, J. R., and Zhang, P. (2013). Functional characterization of Dihydroflavonol-4-reductase in anthocyanin biosynthesis of purple sweet potato underlies the direct evidence of anthocyanins function against abiotic stresses. *PLoS One* 8:e78484. doi: 10.1371/journal.pone.0078484

Wang, Z., Cui, Y., Vainstein, A., Chen, S., and Ma, H. (2017). Regulation of fig (Ficus carica L.) fruit color: Metabolomic and transcriptomic analyses of the flavonoid biosynthetic pathway. *Front. Plant Sci.* 8:1990. eCollection 2017 doi: 10.3389/fpls.2017.01990

Wishart, D. S., Jewison, T., Guo, A. C., Wilson, M., Knox, C., Liu, Y. F., et al. (2013). HMDB 3.0–The Human Metabolome Database in 2013. *Nucleic Acids Res* 41, 801–807. doi: 10.1093/nar/gks1065

Wu, Q., Wu, J., Li, S. S., Zhang, H. J., Feng, C. Y., Yin, D. D., et al. (2016). Transcriptome sequencing and metabolite analysis for revealing the blue flower formation in waterlily. *BMC Genomics* 17:897. doi: 10.1186/s12864-016-3226-9

Wu, X., Gong, Q., Ni, X., Zhou, Y., and Gao, Z. (2017). UFGT: The key enzyme associated with the petals variegation in Japanese apricot. *Front. Plant Sci.* 8:108. doi: 10.3389/fpls.2017.00108

Xie, X. B., Li, S., Zhang, R. F., Zhao, J., Chen, Y. C., Zhao, Q., et al. (2012). The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ* 35, 1884–1897. doi: 10.1111/j.1365-3040.2012.02523.x

Xu, W. J., Grain, D., Bobet, S., Gourrierec, J. L., Thévenin, J., Kelemen, Z., et al. (2014). Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in Arabidopsis seed. *New Phytol.* 202, 132–144. doi: 10.1111/nph.12620

Yang, M. D., Chiang, Y. M., Higashiyama, R., Asahina, K., Mann, D. A., Mann, J., et al. (2012). Rosmarinic acid and baicalin epigenetically derepress peroxisomal proliferator-activated receptorγin hepatic stellate cells for their antifibrotic effect. *Hepatology* 55, 1271–1281. doi: 10.1002/hep.24792

Yao, G. F., Ming, M. L., Allen, A. C., Gu, C., Li, L. T., Wu, X., et al. (2017). Map-based cloning of the pear gene MYB114 identifies an interaction with other transcription factors to coordinately regulate fruit anthocyanin biosynthesis. *Plant J.* 92, 437–451. doi: 10.1111/tpj.13666

Zhang, Q., Wang, L. L., Liu, Z. G., Zhao, Z. H., Zhao, J., Wang, Z. T., et al. (2020). Transcriptome and metabolome profiling unveil the mechanisms of *Ziziphus jujuba Mill.* peel coloration. *Food Chem.* 312, 125903–125910. doi: 10.1016/j.foodchem.2019.125903

Zhao, Q., Chen, X. Y., and Martin, C. (2016b). Scutellaria baicalensis, the golden herb from the garden of Chinese medicinal plants. *Sci. Bull. (Beijing).* 61, 1391–1398. doi: 10.1007/s11434-016-1136-5

Zhao, Q., Zhang, Y., Wang, G., Hill, L., Weng, J. K., Chen, X. Y., et al. (2016a). A specialized flavone biosynthetic pathway has evolved in the medicinal plant. *Scutell. baicalensis. Sci. Adv.* 2:e1501780. doi: 10.1126/sciadv.1501780

Zhou, H., Lin-Wang, K., Wang, H. L., Gu, C., Dare, A. P., Espley, R. V., et al. (2015). Molecular genetics of blood-fleshed peach reveals activation of anthocyanin biosynthesis by NAC transcription factors. *Plant J.* 82, 105–121. doi: 10.1111/tpj.12792

Zhu, Z. J., Schultz, A. W., Wang, J. H., Johnson, C. H., Yannone, S. M., Patti, G. J., et al. (2013). Liquid chromatography quadrupole time-of-flight mass spectrometry characterization of metabolites guided by the METLIN database. *Nat. Protoc.* 8, 451–456. doi: 10.1038/nprot.2013.004

Zhuang, H. M., Lou, Q., Liu, H. F., Han, H. W., Wang, Q., Tang, Z. H., et al. (2019). Differential regulation of anthocyanins in green and purple turnips revealed by combined de novo transcriptome and metabolome analysis. *Int J. Mol. Sci.* 20, 4387–4405. doi: 10.3390/ijms20184387

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 14 June 2022
doi: 10.3389/fpls.2022.923183

Integrative mRNA and Long Noncoding RNA Analysis Reveals the Regulatory Network of Floral Bud Induction in Longan (*Dimocarpus longan* Lour.)

Fan Liang[1†], *Yiyong Zhang*[1†], *Xiaodan Wang*[1], *Shuo Yang*[1], *Ting Fang*[1], *Shaoquan Zheng*[2] *and Lihui Zeng*[1*]

[1] *Insititute of Genetics and Breeding in Horticultural Plants, College of Horticulture, Fujian Agriculture and Forestry University, Fuzhou, China,* [2] *Fujian Breeding Engineering Technology Research Center for Longan & Loquat, Fruit Research Institute, Fujian Academy of Agricultural Sciences, Fuzho, China*

OPEN ACCESS

Edited by:
Shunquan Lin, South China Agricultural University, China

Reviewed by:
Ze Peng, South China Agricultural University, China
Yunxiang Wang, Beijing Academy of Agricultural and Forestry Sciences, China

***Correspondence:**
Lihui Zeng
lhzeng@fafu.edu.cn

[†] *These authors have contributed equally to this work and share first authorship*

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

Received: *19 April 2022*
Accepted: *11 May 2022*
Published: *14 June 2022*

Citation:
Liang F, Zhang Y, Wang X, Yang S, Fang T, Zheng S and Zeng L (2022) Integrative mRNA and Long Noncoding RNA Analysis Reveals the Regulatory Network of Floral Bud Induction in Longan (Dimocarpus longan Lour.). Front. Plant Sci. 13:923183. doi: 10.3389/fpls.2022.923183

Longan (*Dimocarpus longan* Lour.) is a tropical/subtropical fruit tree of significant economic importance. Floral induction is an essential process for longan flowering and plays decisive effects on the longan yield. Due to the instability of flowering, it is necessary to understand the molecular mechanisms of floral induction in longan. In this study, mRNA and long noncoding RNA (lncRNA) transcriptome sequencing were performed using the apical buds of fruiting branches as materials. A total of 7,221 differential expressions of mRNAs (DEmRNAs) and 3,238 differential expressions of lncRNAs (DElncRNAs) were identified, respectively. KEGG enrichment analysis of DEmRNAs highlighted the importance of starch and sucrose metabolic, circadian rhythms, and plant hormone signal transduction pathways during floral induction. Combining the analysis of weighted gene co-expression network (WGCNA) and expression pattern of DEmRNAs in the three pathways, specific transcriptional characteristics at each stage during floral induction and regulatory network involving co-expressed genes were investigated. The results showed that sucrose metabolism and auxin signal transduction may be crucial for the growth and maturity of autumn shoots in September and October (B1-B2 stage); starch and sucrose metabolic, circadian rhythms, and plant hormone signal transduction pathways participated in the regulation of floral bud physiological differentiation together in November and December (B3-B4 stage) and the crosstalk among three pathways was also found. Hub genes in the co-expression network and key DEmRNAs in three pathways were identified. The circadian rhythm genes *FKF1* and *GI* were found to activate *SOC1* gene through the photoperiod core factor *COL* genes, and they were co-expressed with auxin, gibberellin, abscisic acid, ethylene signaling genes, and sucrose biosynthesis genes at B4 stage. A total of 12 hub-DElncRNAs had potential for positively affecting their distant target genes in three putative key pathways, predominantly in a *co*-transcriptional manner. A hypothetical model of regulatory pathways and key genes and lncRNAs during floral bud induction in longan was proposed finally. Our studies will provide valuable clues and information to help elucidate the potential molecular mechanisms of floral initiation in longan and woody fruit trees.

Keywords: longan, transcriptome, mRNA, long noncoding RNA, floral bud induction, WGCNA

INTRODUCTION

Longan (*Dimocarpus longan* Lour.), which belongs to the family of Sapindaceae, is a very important horticultural crop in the tropical and subtropical regions of the world. Floral bud induction largely determines the fruit yield of the next year (Link, 2000). However, the difficulty and instability of blossoming are the most challenging issues in longan production because the apical buds are easily disturbed by various factors during floral induction (You et al., 2012). Therefore, a lot of attention has been paid to the regulatory mechanism of floral induction in longan.

Floral induction is critical in the life cycle of plants, which is influenced by a variety of endogenous and environmental stimuli (Wang and Ke, 1992; Kurokura et al., 2013; Lyons et al., 2015). In the model plant *Arabidopsis*, the molecular mechanisms of flowering have been well established, including vernalization, photoperiod, phytohormones, nutrient responses, and senescence pathways (Wang et al., 2009; Kim et al., 2012; Yamaguchi and Abe, 2012; Cho et al., 2017). Among them, photoperiod and temperature are primary environmental cues. In the photoperiod regulatory pathway, the activity of CONSTANS (CO) protein is tightly controlled by the circadian clock and light signals, thereby affecting integrated genes, such as *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*), to accelerate the flowering process (Pin and Nilsson, 2012). Unlike annual plants, which bloom only once in their life cycle, perennial woody plants can live for many years and flower repeatedly (Albani et al., 2012). However, little is known about the regulatory mechanism of floral bud induction seasonally in woody plants, especially in longan.

Sugars are important endogenous signals in the regulation of plant growth and development, including flowering (Turnbull, 2011). It has been reported that the flowering transition of citrus was regulated by sugar as an energy substance (Shalom et al., 2014). Previous research studies have revealed that T6P, glucose, and sucrose mediated the regulation of floral induction as signaling molecules (Wahl et al., 2013). Sucrose, glucose, and starch are closely related to photoperiod and abscisic acid (ABA) signal and play an indispensable role in floral induction (Su et al., 1997; Curaba et al., 2014; Han et al., 2014). *AtIDD8*, a transcription factor associated with sugar signaling, has been demonstrated to regulate photoperiodic flowering by regulating sugar transport and metabolism in *Arabidopsis* (Seo et al., 2011). The expression of the *CO* gene is affected by the expression of *GBSSI*, a key gene of starch synthase, to promote plant flowering (Serrano et al., 2009).

Phytohormones are also important factors for floral induction. Plant endogenous hormones can regulate plant growth and development by binding to specific protein receptors, and play a micro-efficient regulatory role in the floral bud differentiation of woody plants (Su et al., 1997; Mutasa-Gottgens and Hedden, 2009; Curaba et al., 2014; Han et al., 2014). Gibberellin (GAs) has a positive effect on controlling the onset of flower formation in *Arabidopsis* (Yamaguchi et al., 2014). Physiological and biochemical evidence indicates that auxin and GAs play essential roles in apple floral induction (Xing et al., 2015). Cytokinin (CK) levels have been reported to affect meristem activity and inflorescence complexity in *Arabidopsis* and *Oryza sativa* (Han et al., 2014). Stress hormones, such as abscisic acid (ABA), jasmonic acid (JA), and ethylene (ETH), participate in flowering regulation under abiotic stress through the flowering pathway related to photoperiod and circadian clock (de Montaigu et al., 2010; Shimakawa et al., 2012; Lyons et al., 2015). In fruit trees, hormone's effects on flowering are quite different from model plants, indicating that there may be a more complex regulatory mechanism in fruit trees.

Some flowering genes have been identified in longan. Photoperiodic flowering pathway–related genes *DlGI* and *DlFKF1* were found to induce the flowering of longan by affecting the synthesis and transportation of endogenous auxin (Huang et al., 2017). Both *DlELF4-1* and *DlELF4-2* transgenic *Arabidopsis* plants showed delayed flowering and adventitious root growth, suggesting that they may inhibit flowering and promote auxin synthesis (Fu et al., 2018). These studies suggest that phytohormone and photoperiod may be interrelated in the regulation of flowering in longan.

Another potential regulatory factor in floral induction is long noncoding RNAs (lncRNAs), which are a class of RNA transcripts over 200 nt in length that lack significant protein-coding capacity (Jin et al., 2013). LncRNAs usually have a similar structure to mRNAs. Advancements in lncRNA research have shown that most transcribed regions generate lncRNAs, which may have rich features, not as "transcriptional noise" (i.e., no biological functions). In plants, lncRNAs are involved in diverse biological processes, including flowering and phytohormone-related development (Liu et al., 2015). LncRNAs have been reported in a variety of plant species, including *Arabidopsis thaliana* (Yuan et al., 2016), *Cucumis sativus* (Hao et al., 2015), *Oryza sativa* (Huang et al., 2021), and *Citrus sinensis* (Ke et al., 2019). In *Arabidopsis*, two different classes of lncRNAs (*COOLAIR* and *COLDAIR*) are involved in flowering regulation by repressing *FLOWERING LOCUS C* (*FLC*) *via* epigenetic modulation, respectively (Heo and Sung, 2011; Sun et al., 2013). In rice, *Early flowering-completely dominant* (*Ef-cd*) encoding an antisense lncRNAOsSOC1 overlapping a floral activator shortens maturity without concomitant yield loss (Fang et al., 2019). It is suggested that flowering genes and related lncRNAs may play a joint role in floral induction and flowering in plants. However, little is known about lncRNAs and their potential roles during floral induction in fruit trees.

The purpose of this study is to explore the transcriptional regulation of mRNA and lncRNA during floral induction in longan. DEmRNAs and DElncRNAs related to floral bud initiation were identified, and the *cis*- and *trans*-target genes of DElncRNAs were systematically predicted and analyzed. The potential relationship between mRNA and lncRNA during longan floral induction was analyzed. Furthermore, regulatory networks and key candidate genes in response to longan floral bud induction were identified by constructing co-expression networks. Our results will reveal novel insights into the regulatory mechanisms underlying seasonal flowering induction in longan and provide strategies for increasing longan yield.

MATERIALS AND METHODS

Plant Materials

In the current study, apical buds of fruiting branches (diameter over 0.8 cm) of "Lidongben" (subsequently referred to as "LDB") longan from September 2019 to January 2020 were used as experimental materials. The mature "LDB" trees grow in the experimental field of Fruit Tree Research Institute of Fujian Academy of Agricultural Sciences. The apical buds of five stages (B1, September; B2, October; B3, November; B4, December; B5, January) were collected. Samples were collected from 10 am to 12 am on the 20th of each month and each stage had three biological replicates. After collection, the buds were immediately frozen in liquid nitrogen and stored at −80°C until used. According to China Weather Network (http://www.weather.com.cn/jiangsu), the climate data of Fuzhou from September 2019 to January 2020 was recorded in **Supplementary Table 1**.

Library Preparation for Transcriptome Sequencing and Identification of mRNA and lncRNA

A total amount of 1 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext UltraTM RNA Library Prep Kit for Illumina (NEB, United States) following the manufacturer's recommendations. The library preparations were sequenced on an Illumina platform and paired-end reads were generated. All clean reads generated in this study were deposited in the NCBI Sequence Read Archive database (http://www.ncbi.nlm.nih.gov/sra/) under the project accession number PRJNA830603.

Annotated mRNAs were obtained by comparison with CPC2[1] / PFAM[2] / CNCI[3] databases using Cuffcompare software. Cuffmerge software was used to merge the transcripts obtained by the splicing of all samples, and removed the transcripts with uncertain chain directions and transcripts with a length of less than 200 nt. Next, Cuffcompare software was used to screen out transcripts that overlapped with the exon region annotated in the database, and included the lncRNAs in the database that overlapped with the exon region of the spliced transcripts as database annotation lncRNAs in the subsequent analysis. The transcripts without coding potential were lncRNAs. For the transcripts obtained in the previous step, the current mainstream coding potential analysis databases (CPC2 / PFAM / CNCI) were combined to predict the coding potential, and the intersection of the transcripts with coding potential in these software results was taken as the Novel_mRNA predicted by this analysis. Take the intersection of transcripts without coding potential in these software results as the candidate Novel_lncRNA data set predicted by this analysis (Harrow et al., 2012). According to the naming guidelines of HGNC[4] (Povey et al., 2001), the candidate novel_lncRNA was finally screened and named according to its positional relationship with coding genes. The lncRNA sequences of *Malus domestica*, *Citrus sinensis*, and *Arabidopsis thaliana* were obtained from the Green Non-coding Database[5] (GREENC) (Paytuvi Gallart et al., 2016) for sequence conservation analysis. Conservation analysis of lncRNAs was performed using the BLAST tool of integrative toolkit TBtools (Chen et al., 2020).

Analysis of DEmRNAs and DElncRNAs

Fragments per kilobase million (FPKM) were used to quantify mRNAs and lncRNAs expression levels in each sample using StringTie software. Screening and identification of DEmRNAs and DElncRNAs using edgeR software (Robinson et al., 2010). The DEmRNAs and DElncRNAs were identified based on the following parameters: $|\log_2\text{foldchange}| \geq 1$ and $p < 0.05$. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed using DEmRNAs that were differentially combined in different stages.

Target Gene Prediction of lncRNA

There are various mechanisms by which lncRNAs regulate target genes, and the most common ways to target genes are *cis* and *trans* target gene regulation. According to the positional relationship between lncRNA and mRNA, the protein-coding gene transcribed within 100 kb upstream and downstream of the lncRNA was identified as the *cis* target gene (Bao et al., 2019). In addition, *trans* target genes were determined by calculating the correlation between lncRNAs and mRNAs expression levels, and the correlation of expression was assessed using Pearson's correlation coefficient ($r > 0.95$) (Kopp and Mendell, 2018). The $\log_2$transformed (FPKM + 1) value was used to normalize the expression levels of mRNA and lncRNA. The expression levels of DEmRNAs and DElncRNAs were heatmapped with integrative toolkit TBtools on the basis of normalized FPKM values. LncRNA sequences were used as queries to discover the enriched motifs using online server STREME[6] (Bailey, 2021) ($p < 0.05$). The possible roles of each identified motif was predicted using online server GOMO (gene ontology for motifs[7]) (Buske and Bodén, 2010). In the MEyellow and MEblue modules, transcription factors (TFs) were identified and classified according to the reference Plant TFDB 4.0 database[8].

Coexpression Network Analysis of DEmRNAs and DElncRNAs

R package WGCNA (Zhang and Horvath, 2005; Langfelder and Horvath, 2008) was used to test whether the expression levels of DEmRNAs and DElncRNAs were correlated with floral bud induction and the association between DEmRNAs and DElncRNAs. In this analysis, the soft threshold power was set to 5, and then an adjacency matrix was constructed using the adjacency function. Based on the transformation of the adjacency matrix into a topological overlap metric graph, similarity matrices of DEmRNAs and DElncRNAs expression between different nodes were calculated. Based on this algorithm, DEmRNAs and DElncRNAs were hierarchically clustered and

[1]https://cpc.cbi.pku.edu.cn/
[2]http://pfam.xfam.org/
[3]https://www.bioinfo.org/software/cnci/
[4]https://www.genenames.org/
[5]http://greenc.sequentiabiotech.com/wiki/Main_Page
[6]https://meme-suite.org/meme/tools/streme
[7]https://meme-suite.org/meme/tools/gomo
[8]http://planttfdb.gao-lab.org/prediction.php

modules were defined. As the method described by Gerttula et al. (2015), the modules were associated with stages and the correlation matrix of both was computed. The module with Pearson's correlation coefficient >0.75 or < −0.5 and $p < 0.05$ were designated as the module most correlated with the stages. The module membership (MM) value of the co-expression network for each module was calculated and used to rank members in each module and identify hub-genes according to MM > 0.75. On this basis, the network of co-expression centers was analyzed and mapped by Cytoscape (version 3.7.1) (Jiang et al., 2017).

Expression Validation of Candidate Gene by qRT-PCR

Total RNA was isolated and extracted from longan apical buds using RNAprep Pure Plant Plus Kit (TIANGEN, China), and was reverse transcribed into cDNA with *5×TransScript* Uni All-in-One SuperMix and gDNA Remover in the *PerfectStart* Uni RT&qPCR Kit (Transgen, China) according to the manufacturers' instructions. qRT-PCR was performed on the instrument of *LightCycler 96* PCR System (Roche, Switzerland) using *PerfectStart* Green qPCR SuperMix (Transgen,China). The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression levels of candidate genes (Jue et al., 2018). The actin gene *DlACT* (**Supplementary Table 2**) was used as the internal reference to normalize expression levels. The primer sequence information was shown in **Supplementary Table 2**. Three biological replicates were used for each experiment.

RESULTS

Development of Apical bud During Floral Induction in Longan

In Fujian Province, autumn shoots of longan grow and get maturity in September (B1 stage) and October (B2 stage). Then in November and December (B3, B4 stage), apical buds of fruiting branches turn into the stage of floral bud physiological differentiation, followed by morphological differentiation in January (B5 stage) (Wang and Ke, 1992; Yuan, 2016). The phenotypes of apical buds of "LDB" longan at B1, B2, B3, B4, and B5 stages were shown in **Figure 1**. The apical buds at B1 (**Figure 1A**) and B2 (**Figure 1B**) stages were shorter overall compared to other stages. The axis of the apical bud is elongated and axillary buds can be seen in the axil at B3 and B4 stages (**Figures 1C,D**). At the B5 stage, the apical buds turned red (called "red bud") (**Figure 1E**).

Characterization of mRNA and lncRNA

Through genomic comparison, CPC, CNCI, and PFAM database analysis, 39,281 annotated mRNAs, 1,780 novel mRNAs (**Figure 2A**), and 34,702 putative lncRNAs (**Figure 2B**) were identified (**Supplementary Table 3**). The most abundant class was seen overlapping lncRNAs (80%), followed by antisense lncRNAs (16%) and lncRNAs (4%) (**Figure 2C**) in putative lncRNAs. FPKM was used to estimate the expression levels of mRNAs and lncRNAs (**Figure 2D**), and the results revealed that the overall expression level of mRNAs was higher than that of lncRNAs, and the expression level of most lncRNAs was low (FPKM ≤ 1). Moreover, the basic characteristics of mRNAs and lncRNAs, including transcript length, ORF length, and number of exons, were analyzed. The results showed that the total length and ORF length of mRNAs were longer than that of lncRNAs (**Figures 2E,F**), and the number of exons of mRNAs is more than that of lncRNAs (**Figure 2G**). In addition, significant differences were detected in the distribution of exons of mRNAs and lncRNAs. A 91.5% of lncRNAs had only one or two exons, while mRNAs had more exons and wider distribution. These results verified that the predicted mRNAs and lncRNAs were in line with general characteristics.

Sequence similarity and conservation are regarded as indicators of biological function. The conservative analysis of lncRNAs was assessed by using E-value < 1e-5. In longan, 28, 19, and 78 lncRNAs were sequentially conserved with *Arabidopsis thaliana* (15), *Malus domestica* (12), and *Citrus sinensis* (79), respectively (**Supplementary Table 4**), indicating that lncRNAs were less conserved among different species.

Differential Expression and Functional Analysis of mRNA and lncRNA

A total of 7,221 DEmRNAs were identified by time-series pairwise comparisons, which included 1,677 DEmRNAs in B2_*vs*_B1, 1,139 in B3_*vs*_B2, 1,832 in B4_*vs*_B3, and 3,961 in B5_*vs*_B4 (**Figure 3A**; **Supplementary Table 5**). A total of 3,238 DElncRNAs were identified, including 1,103 in B2_*vs*_B1, 950 in B3_*vs*_B2, 1,244 in B4_*vs*_B3, and 1,021 in B5_*vs*_B4 (**Figures 3B,D**; **Supplementary Table 5**). The transition from B4 stage to B5 stage resulted in more complex transcriptional changes in DEmRNAs than that in other stages.

A lncRNA/protein-coding gene regulatory pair consists of a controlled lncRNA and its regulated protein-coding gene. LncRNA has been reported to regulate the expression of protein-coding genes in both *cis*- and *trans*-regulatory mechanisms (Hu et al., 2016; Wei et al., 2019). On the basis of 3,238 DElncRNAs as studied, a total of 4,124 *cis*-regulated lncRNA-mRNA pairs were identified in the 100 kb range, including 1,947 DElncRNAs and 2,007 DEmRNAs (**Supplementary Table 6**). For lncRNA *trans*-targets, total of 13,840 regulatory pairs were identified, including 1,069 DElncRNAs and 1,482 DEmRNAs (**Supplementary Table 6**). These findings indicate that lncRNAs may demonstrate rich functions in *trans*-regulating their distant genes and co-location regulation.

Pathway-based analysis helps to further understand the biological function of genes. To explore transcriptional changes during floral induction, KEGG enrichment analysis was performed on DEmRNAs at different stages. Abundant metabolic and signaling pathways were enriched in all comparison groups, and pathways implicated in flowering, such as "starch and sucrose metabolism," "circadian rhythms – plant," and "plant hormone signal transduction", were found. KEGG analysis showed that "starch and sucrose metabolism" was significantly enriched in the top 20 enriched pathways in all comparison groups. The "plant hormone signal transduction" pathway was significantly enriched in B2_*vs*_B1 and B4_*vs*_B3

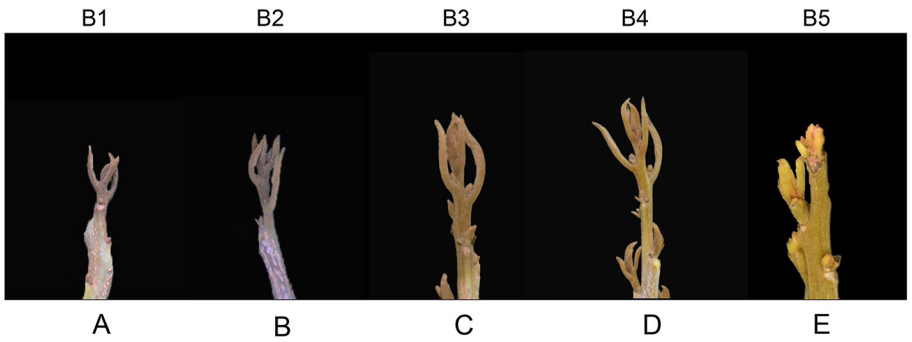


FIGURE 1 | Morphological characteristics of the apical bud during floral bud induction in longan. **(A–E)** Morphological characteristics of apical buds at B1, B2, B3, B4, and B5 stages, respectively.

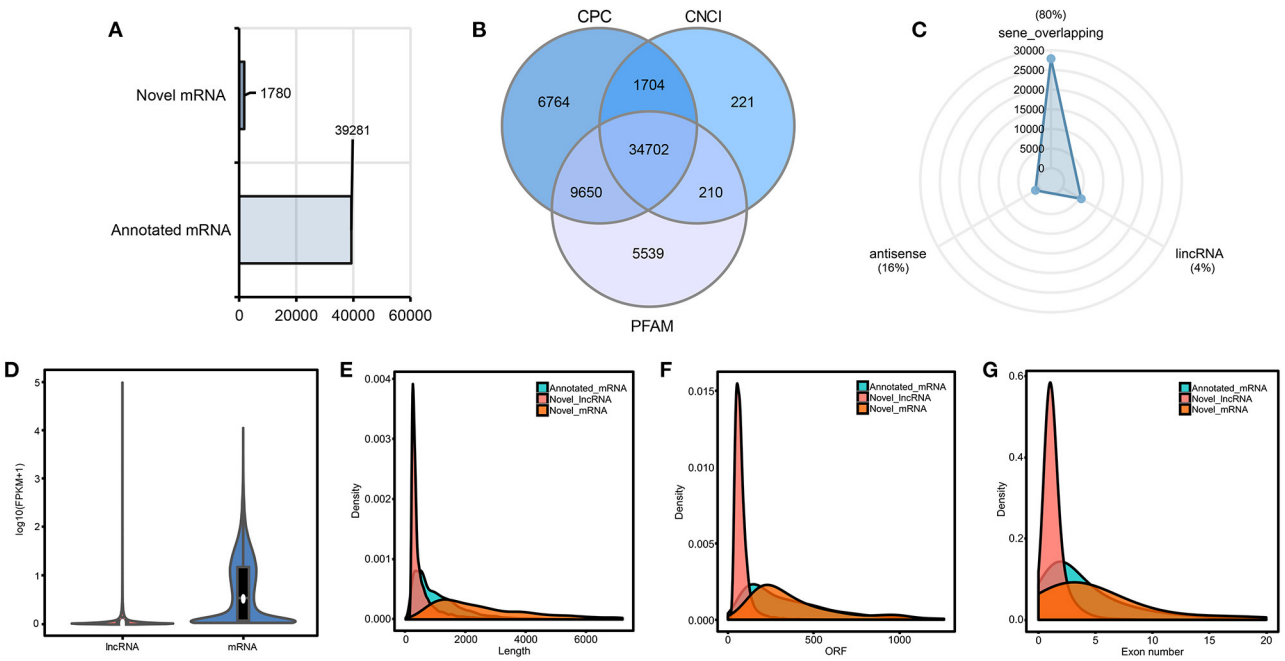


FIGURE 2 | Characterization and statistical analysis of candidate mRNAs and lncRNAs. **(A)** Statistics of novel mRNA and annotated mRNA. The horizontal axis indicates the number of mRNAs. **(B)** Venn diagram of candidate lncRNAs. **(C)** Classification of the identified lncRNAs. **(D)** FPKM distribution of mRNA and lncRNA transcripts. **(E–G)** Characterization of annotated mRNA, novel mRNA, and novel lncRNA transcripts. The horizontal axis of the density plot from left to right indicates total length (left), ORF length (middle), and the number of exons (right) of annotated mRNA, novel mRNA, and novel lncRNA transcripts, respectively.

(**Figures 3C,E**). In addition, the "circadian rhythm–plant" pathway was significantly enriched in B4_*vs*_B3 and B5_*vs*_B4 (**Figures 3E,F**). Three flowering-related KEGG pathways were also identified among the target genes of the DElncRNAs (**Supplementary Table 6**). The results showed that these three biological processes were transcriptionally active during floral induction in longan.

Co-Expression Networks Reveal Modules Specific to Longan Floral Induction

To understand the dynamic changes of represented genes and systematically explore the potential regulatory function of lncRNA associated with longan floral bud differentiation, WGCNA was performed to analyze the co-expression relationship between 7,221 DEmRNAs and 3,228 DElncRNAs. Module–stage association analyses were performed to select stage-specific modules. A total of 17 different modules were obtained. MEblack, MEdarkred, MEyellow, MEblue, and MEpurple modules were highly correlated with different stages of longan floral induction (Pearson's correlation coefficient $>$ 0.75 or $<$ -0.5 and $p < 0.05$) (**Figure 4A**). KEGG enrichment (**Figure 4B**) and major expression center analyses were also conducted for the five modules. "Plant hormone signal transduction" was the most abundant pathway among the

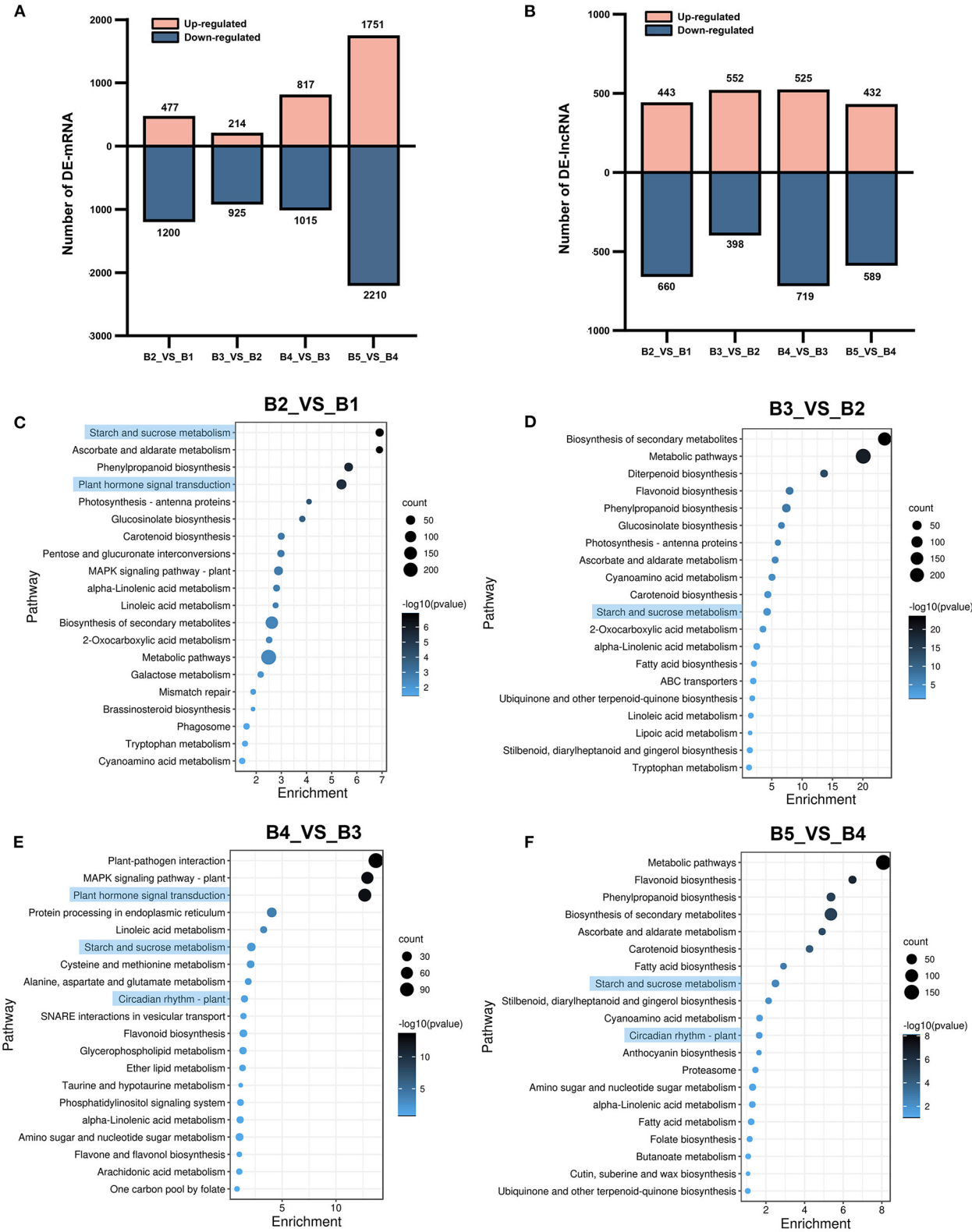


FIGURE 3 | Up- and down-regulated numbers of DEmRNAs and DElncRNAs and KEGG enrichment analysis of DEmRNAs at different stages. **(A)** The number of up-regulated and down-regulated DEmRNAs. **(B)** The number of up-regulated and down-regulated DElncRNAs. **(C–F)** KEGG enrichment analysis of DEmRNAs in B2_*vs*_B1, B3_*vs*_B2, B4_*vs*_B3, and B5_*vs*_B4. The horizontal axis shows -$\log_{10}(p)$. The vertical axis shows pathways of enrichment.

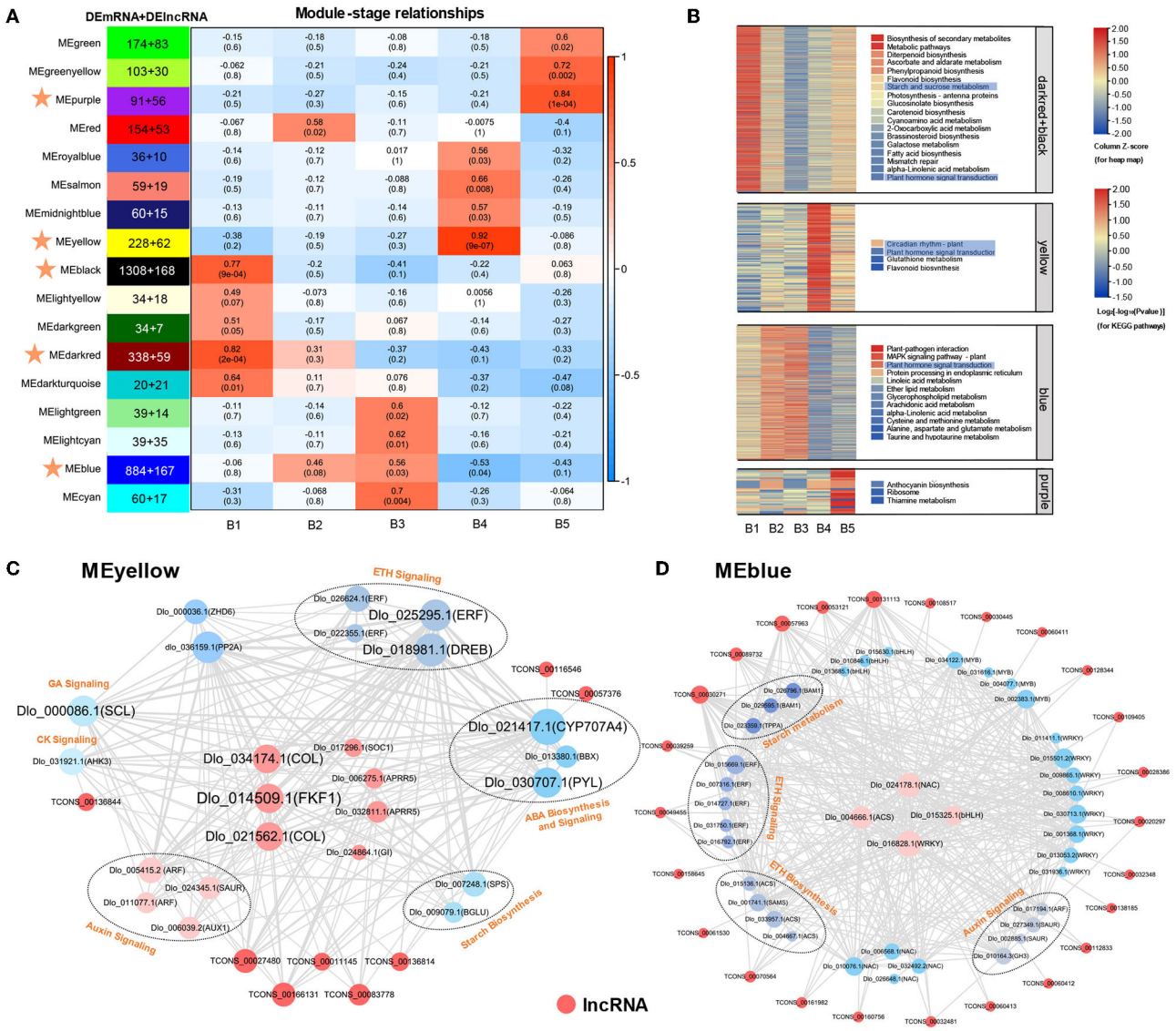


FIGURE 4 | Co-expression network analysis of DEmRNAs and DElncRNAs. **(A)** Heatmaps showing correlation of module eigengenes at different stages of floral induction in longan. Pearson correlation coefficient is given in each module for the different stages and is colored from blue to red according to the scores. The MEpurple, MEyellow, MEblack, Medarkred, and MEblue modules (Pearson's correlation coefficient >0.75 or <-0.5 and $p < 0.05$) with stars are closely related to the floral induction in longan. **(B)** KEGG enrichment analysis and transcriptional changes of DEmRNAs in the MEpurple, MEyellow, MEblack, MEdarkred, and MEblue modules. The selected module KEGG enriched pathways are visualized by different colors. **(C,D)** Visualization of the central coexpression network in MEyellow and MEblue modules associated with B4 stage. The size of nodes indicates connectivity. The width of the line indicates the strength of the correlation.

five modules. Closer inspection revealed that DEmRNAs in MEyellow and MEblue modules, which were highest associated with B4 stage, were significantly enriched in "circadian rhythm" and "plant hormone signal transduction," suggesting that phytohormone and circadian rhythm may play a dominant role in B4 stage. The MEblack and MEdarkred modules were highest correlated with the B1 stage, and genes of "starch and sucrose metabolism" and "plant hormone signal transduction" pathways were significantly enriched in these two modules, indicating that sucrose and phytohormone played an important role in the regulatory network of B1 stage. KEGG enrichment analysis of DEmRNAs in these highly correlated 5 modules also showed the interaction between phytohormone signaling and circadian rhythms at the B4 stage, and the interaction between starch and sucrose metabolism and phytohormone signaling at the B1 stage (**Figure 4B**). The results of WGCNA were consistent with the results of KEGG enrichment analysis, suggesting that three regulatory pathways of starch and sucrose metabolism, phytohormone signaling and circadian rhythm may play important roles during longan floral induction.

Module membership (MM) represents the degree of connection based on the eigengene, which is a measure of a

gene relative to the membership of the module and is used to screen the hub gene. DEmRNAs and DElncRNAs with MM > 0.75 were identified as hub genes. A total of 1,976 hub-DEmRNAs and 207 hub-DElncRNAs were identified (**Supplementary Table 7**). For example, highly connected hub genes in MEyellow module included "*two-component response regulator 5* (*APRR5*)," "*CONSTANS-LIKE* (*COL*)," "*FKF1,*" "*GI,*" and "*SOC1,*" which were well-known genes in circadian rhythm and photoperiod regulatory pathway during flowering. The central expression networks of MEyellow and MEred modules at the B4 stage were further analyzed (**Figures 4C,D**). The main hubs in the co-expression network of MEyellow module included rhythm-related genes (*COL*, *Dlo_034174.1*, *Dlo_021562.1*; *FKF1*), *ethylene-responsive transcription factor* (*ERF*, *Dlo_025295.1*), and ABA-degrading gene *CYP707A4* (**Figure 4C**), which were co-expressed with auxin, gibberellin, ethylene signaling, and sucrose biosynthesis genes, which were highly positively correlated. MEblue module contained more transcription factors, such as *bHLH*, *WRKY*, *MYB,* and *NAC*, which were highly negatively correlated, among them, *NAC* (*Dlo_024178.1*), *WRKY* (*Dlo_016828.1*), *bHLH* (*Dlo_015325.1*), and *ACS* (*Dlo_004666.1*) had the highest connectivity (**Figure 4D**). Similarly, MEblue hub genes were co-expressed with genes involved in phytohormone biosynthesis, signal transduction, and sucrose metabolism. Our findings revealed the crosstalk of three regulatory pathways during longan floral induction.

Interaction patterns between lncRNAs and mRNAs were also recognized within MEyellow and MEblue networks (**Figures 4C,D**). There were multiple interaction patterns between lncRNAs and mRNAs, such as one-to-one, one-to-many, and many-to-many. Hub-DElncRNAs are associated with multiple regulatory pathways. For example, TCONS_00166131 has the potential to regulate the photoperiod-related *COL* (*Dlo_034174.1*, *Dlo_021562.1*) and *FKF1* genes, the ethylene signaling-related *ERF* gene, and the sucrose biosynthesis-related *SPS* gene within the same module (**Figure 4C**), revealing that DElncRNAs may affect the expression of targeting genes in different pathways and serve as the regulator of gene co-expression networks. Transcripts, including DElncRNAs and DEmRNAs, in the same module have the same expression pattern and may also share related functions or biological pathways. Meanwhile, the results also showed that DElncRNAs had the potential for *trans*-regulating a spatial number of genes within modules, as opposed to tightly restricted *cis*-regulation.

DEmRNAs in Sucrose Metabolism and Signaling Pathway

A total of 30 DEmRNAs associated with carbohydrate biosynthesis and metabolism were identified. In sucrose biosynthesis, the expression levels of most DEmRNAs were up-regulated at the B1 stage, especially *sucrose synthase* (*SUS*) and *trehalose-phosphate synthase* (*TPS*). For starch degradation, a stronger up-regulation of DEmRNA expression was found at B3 and B4 stages, such as *beta-amylase* (*BAM*) and *4-α-glucanotransferase* (*DPE*) (**Figure 5A**). The transcript level of *sucrose phosphate synthase* (*SPS*) gene, which was one of the key enzymes for sucrose entry into metabolic pathways, was up-regulated at B4 stage. Most of the DEmRNAs involved in sucrose biosynthesis and metabolism were found in MEblack and MEdarkred modules, the majority of which included genes encoding *TPS*, *β-glucosidase* (*BGLU*, *BGLX*), *hexokinase* (*HXK*), and others (**Figure 5C**).

DEmRNAs in Photoperiod and Circadian Clock Pathway

The circadian pathway was one of the main enriched pathways by KEGG analysis. A total of 18 DEmRNAs were found in this pathway, most of them were unique to the MEyellow module and the expression levels were significantly up-regulated at the B4 stage (**Figure 6A**). Three *COL* genes were found located downstream of *FKF1* and *GI* genes, which had the same expression trend as *FKF1* and *GI*. The same up-regulated expression trend was also found for two orthologs of *Pseudo-response Regulator 5* (*PRR5*, *Dlo_006275.1*, *Dlo_032811.1*), as well as *SOC1*and *TCP*. A potential circadian regulatory pathway was constructed during floral induction in longan based on the identified DEmRNAs.

DEmRNAs in Phytohormone Biosynthesis and Signal Transduction

Analysis of genes in hormone signaling pathways revealed at least two-fold significant changes in expression levels between different stages of floral induction. 29 and 10 DEmRNAs were identified in auxin and GA biosynthesis and signaling pathways, respectively, and 19 DEmRNAs were associated with ETH biosynthesis and signal transduction (**Figure 7A**). In the auxin pathway, two *YUCCA* genes, which positively regulate auxin biosynthesis, had opposite expression patterns. The expression of multiple *small auxin up RNAs* (*SAURs*) and *Aux/IAA* genes were down-regulated at the B3 stage. The expression of *ARF* (*Dlo_017194.1*, *Dlo_011077.1*, *Dlo_005415.2*) was up-regulated at B3 or B4 stage. In the GA biosynthesis pathway, the expression levels of two *GA20OX* genes were down-regulated at B3 and B4 stages, implying a decrease in the rate of gibberellin synthesis. In the ETH pathway, the rate-limiting enzymes of ethylene biosynthesis, namely, *1-aminocyclopropane-1-carboxylate oxidase* (*ACO*) and *1-aminocyclopropane-1-carboxylate synthase* (*ACS*), were both down-regulated at the B4 stage. The lower transcriptional levels of seven ethylene biosynthesis-related genes suggested that the ethylene accumulation may be reduced at B4 stage. Members of *ERF* gene family showed different expression patterns, among which three *ERF* genes, namely, *Dlo_009939.1*, *Dlo_007316.1*, *Dlo_015669.1*, had the same expression trend as biosynthetic genes.

Potential Hub-DElncRNAs With Their Target Genes in Three key Pathways

To determine the role of 229 Hub-DElncRNAs during longan floral induction, their co-localized and co-expressed protein-coding genes were predicted; 33 hub-DElncRNAs were related

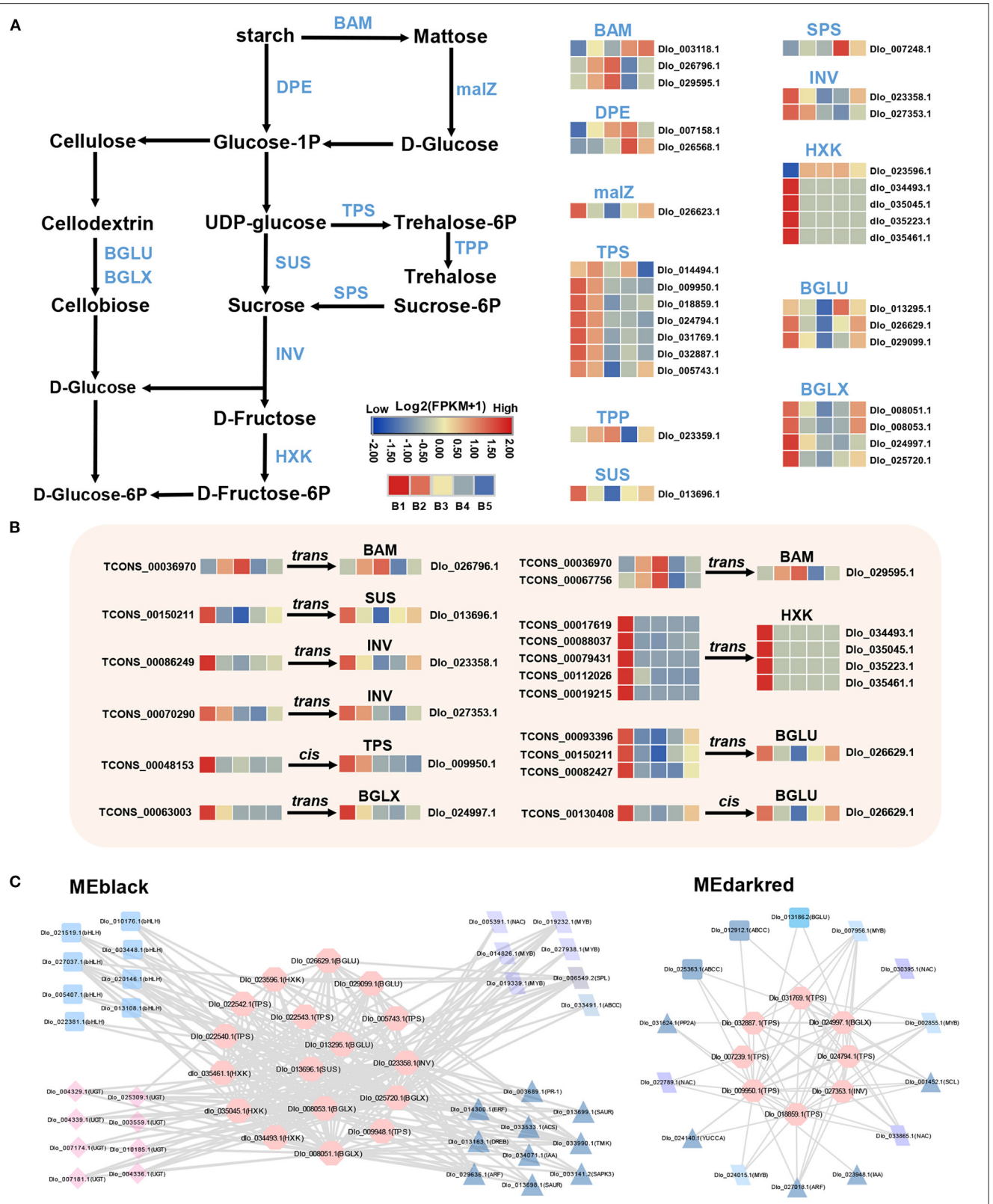


FIGURE 5 | DEmRNAs and DElncRNAs in sucrose metabolism and signaling pathway. **(A)** Transcriptional changes in DEmRNAs associated with sucrose metabolism. **(B)** Expression profiles of hub-DElncRNAs and their target genes related to the sucrose metabolism based on the transcriptome data. Transcriptional change levels were normalized by $\log_2$transformed (FPKM + 1). **(C)** Co-expression network of genes involved in sucrose biosynthesis and metabolism in MEblack and Medarkred (pink nodes). Different nodes represent different types of genes, and edges connect co-expression genes.

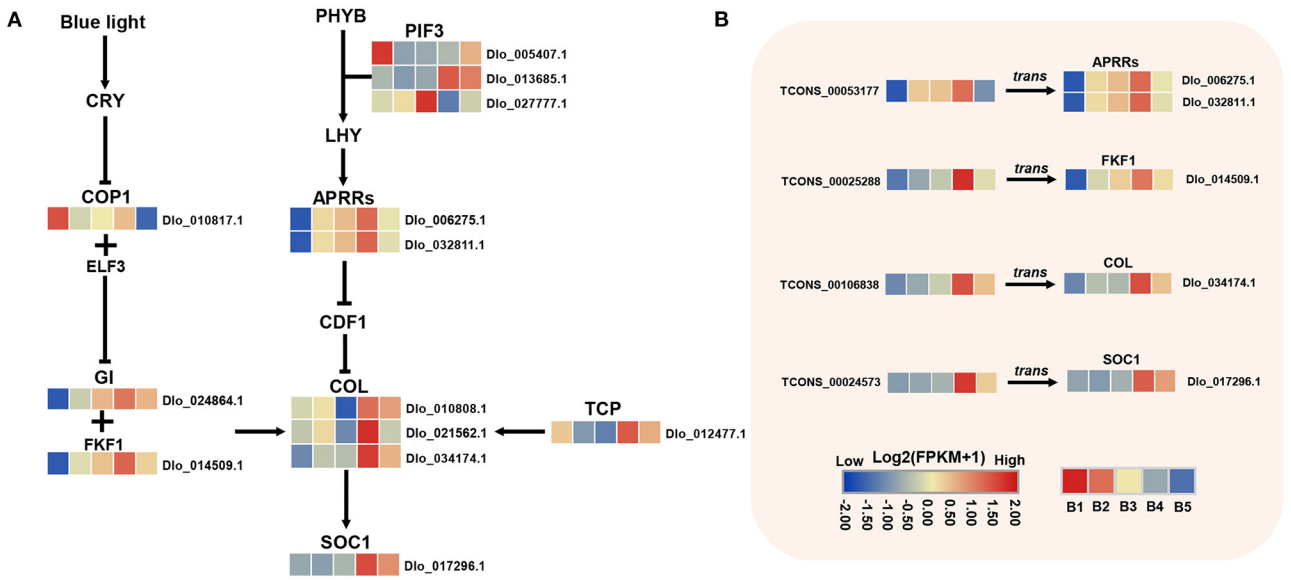


FIGURE 6 | DEmRNAs and DElncRNAs in photoperiod and circadian clock pathways. **(A)** Transcriptional trends of DEmRNAs in photoperiod and circadian clock pathway and visualized as heatmaps. **(B)** DEmRNAs targeted by hub-DElncRNAs. Expression profiles of hub-DElncRNAs and their target genes related to photoperiod and circadian clock based on the transcriptome data. Transcriptional change levels were normalized by log_2transformed (FPKM + 1).

to sucrose metabolism, circadian rhythm, and phytohormone biosynthesis and signaling pathways, and the specific lncRNA-mRNA regulatory relationship was shown in **Figures 5B**, **6B**, **7B**; **Supplementary Table 8**.

Correlation analysis of neighboring lncRNA–mRNA pairs (**Supplementary Table 9**) revealed that three Hub-DElncRNAs (TCONS_00048153, TCONS_00130408, TCONS_00053177) and their neighboring DEmRNAs (*TPS*, *Dlo_009950.1*; *BGLU*, *Dlo_026629.1*; *ARF*, *Dlo_011077.1*) were positively correlated (Pearson's correlation coefficient $\geq$ 0.58, $p \leq 0.05$), and one neighboring lncRNA-mRNA pair had a negative correlation, suggesting that hub-DElncRNAs regulated neighboring protein-coding genes *via cis*-regulatory manner, mainly in a positive manner. In addition, neighboring lncRNA–mRNA pairs may be activated by the same promoter or enhancer for simultaneous transcriptional expression (Ye et al., 2022). To investigate the specific regulatory mechanism of the positive correlation, random lncRNA–mRNA pairs were used as the comparison group for analyzing. The results showed that the correlations of the three neighboring lncRNA–mRNA pairs were statistically significant compared with random lncRNA–mRNA pairs (p=0.012) (**Figure 8A**), indicating that neighboring lncRNA–mRNAs shared the same promoters or enhancers partly.

To further elucidate the *trans*-regulatory roles of lncRNAs in the three pathways, a total of 31 hub-DElncRNAs with *trans*-regulatory potential were identified, among them, there were 16 sucrose metabolism-related, 5 circadian rhythm-related, and 29 hormone signal transduction-related lncRNA-mRNA pairs. Correlation analysis showed that these lncRNA–mRNA pairs had positive correlations, suggesting that hub-DElncRNAs have positive effects on these distant genes.

In the detection of the regulatory relationship between lncRNA and mRNA, 15 hub-DElncRNAs were found to target multiple genes, and 16 hub-DElncRNAs had only one target gene. Sucrose metabolism-related genes (*BAM*, *Dlo_026796.1*; *SUS*, *Dlo_013696.1*, and others), four photoperiod-related genes (*FKF1*, *APRR5*, *COL*, and *SOC1*), and four hormone signal transduction-related genes (*AUX1*, *Dlo_006039.2*; *ARF*, *Dlo_005415.2*; *ERF*, *Dlo_025295.1*, *Dlo_015669.1*) were target genes and they were regulated by a single hub-DElncRNA, respectively. One-to-many and many-to-many regulatory relationships between hub-DElncRNAs and potential target genes were also observed. For example, TCONS_00025288 had three targets, the rhythm gene *FKF1*, *ARF* (*Dlo_005415.2*), and *ERF* (*Dlo_025295.1*). These results suggested that lncRNAs may act as the regulator in different pathways and play an important role in the regulatory network of floral induction.

LncRNAs have also been reported to regulate gene expression by binding to transcription factors (TFs) directly (Long et al., 2017; Ye et al., 2022). Using the MEyellow and MEblue modules as examples for predicting intra-module transcription factors, several highly connected TFs (including *bHLH*, *WRKY*, *ERF*, and *NAC*) were found were co-expressed with sucrose metabolism, photoperiod, and hormone-related genes in the co-expression network (**Figure 8B**). LncRNA has the possibility to regulate 3 putative pathway-related genes by binding to these TFs. Therefore, the identification of binding motifs in lncRNAs will help predict the interactions between lncRNAs and their target TFs. "AAAAAAAAGAAA," "TAATTAAAA," and "TAAHTTTA" were identified as three major binding motifs of lncRNA sequences (**Figures 8C,E,G**; **Supplementary Table 10**). GOMO analysis (**Figures 8D,F,H**; **Supplementary Table 10**) showed that

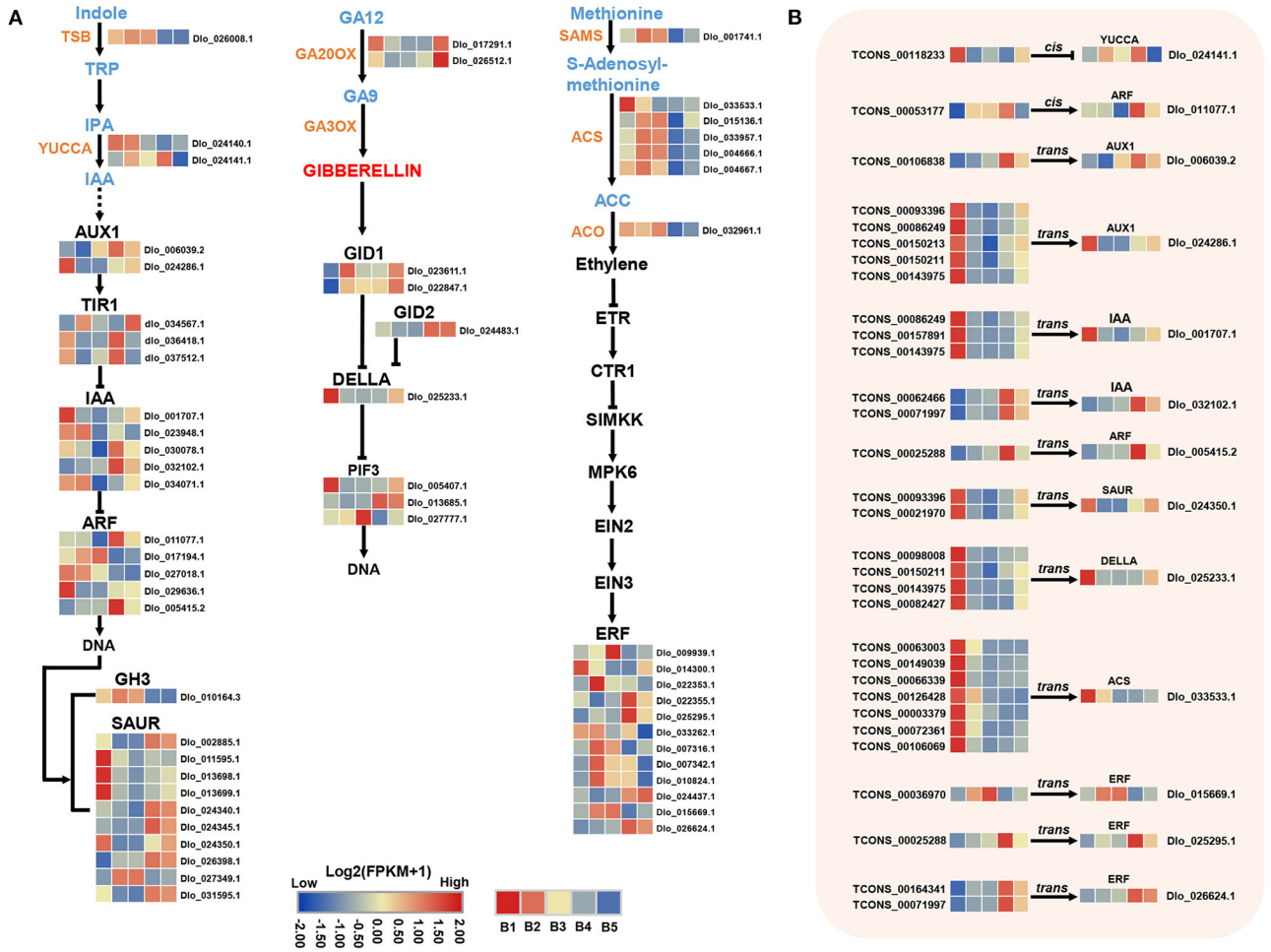


FIGURE 7 | DEmRNAs and DElncRNAs in auxin, gibberellin (GAs), and ethylene (ETH) biosynthesis and signal transduction pathways. **(A)** Transcriptional changes of genes associated with auxin, GAs, and ETH at different stages visualized as heatmaps. **(B)** Expression profiles of hub-DElncRNAs and their target genes related to the phytohormone biosynthesis and signal transduction based on the transcriptome data. Transcriptional change levels were normalized by $\log_2$transformed (FPKM + 1).

the three motifs had transcription factor activity, DNA binding, and transcriptional regulatory functions, indicating that lncRNAs had great potential to interact with TFs through both transcriptional and post-transcriptional regulatory ways. The TF families co-expressed with lncRNAs in modules are worthy of further study.

qRT-PCR Validation of DEmRNAs and DElncRNAs

The transcript abundances of 12 DEmRNAs and 5 DElncRNAs involved in three presumptive pathways were analyzed by qRT-PCR (**Figures 9A,B**; **Supplementary Table 11**). The qRT-PCR results showed that expression levels of *ARF*, *IAA*, *AHK*, *COL*, *FKF1*, *GI*, and *SCL* were significantly up-regulated at the B4 stage and were consistent with the transcriptome data. The same variation trends for the DEmRNAs were well shown between qRT-PCR and FPKM values, indicating the reliability of the RNA-Seq data. The expression levels of 5 DElncRNAs were also analyzed by qRT-PCR. The expression levels of DElncRNAs were consistent with their target genes. For example, the expression level of TCONS_00025288 was consistent with its target gene *FKF1*, showing a positive correlation.

DISCUSSION

Three Important Regulatory Pathways During Floral bud Induction of Longan

In this study, transcriptome sequencing was performed to study the transcriptional changes and regulatory pathways during floral induction in longan. From the results of KEGG enrichment and WGCNA analysis of DEmRNAs, three pathways ("starch and sucrose metabolism," "circadian rhythm," and "plant hormone signal transduction") were found to play an important role during floral bud induction. At B1 and B2 stages, vigorous sucrose metabolism was observed, and phytohormone signal

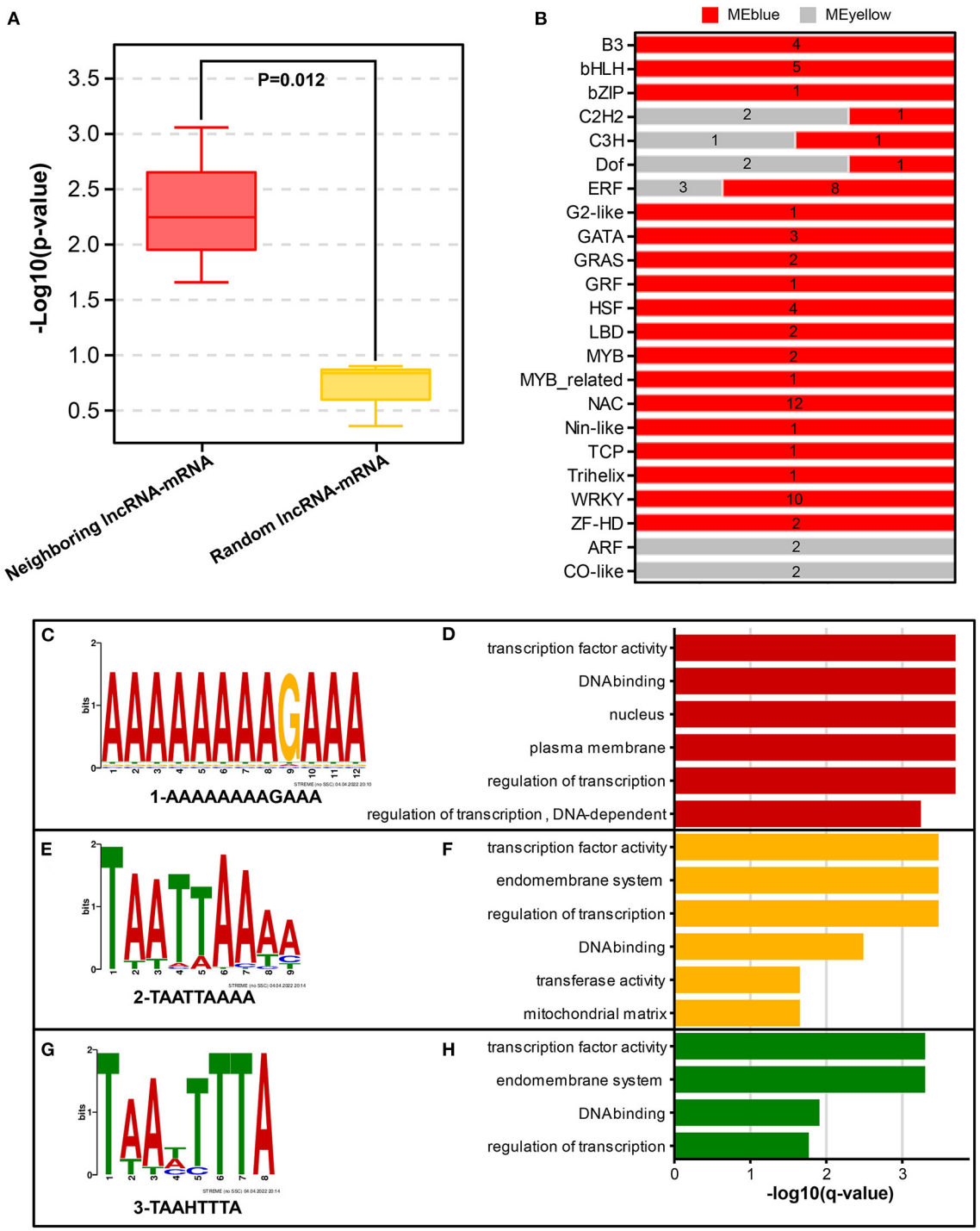


FIGURE 8 | Potential regulatory modes of lncRNAs. **(A)** Comparison of co-expression correlations between neighboring lncRNA-mRNA and random lncRNA-mRNA groups. Significant differences between the two groups were analyzed using Duncan's test. **(B)** Quantity distribution of transcription factors (TFs) included in two example modules (MEblue and MEyellow). **(C,E,G)** Three representative binding motifs in lncRNA sequences from MEyellow and MEblue modules. **(D,F,H)** Functional analysis of 3 representative lncRNA-binding motifs using GOMO tool.

transduction was also found to play an important role in this period. In Fujian Province, B1 and B2 stages (September and October) are the time when autumn shoots grow and get maturity (Yuan, 2016). The level of endogenous nutrient accumulation is directly related to the development of autumn shoots in longan (Li et al., 2013). Therefore, our results confirm that

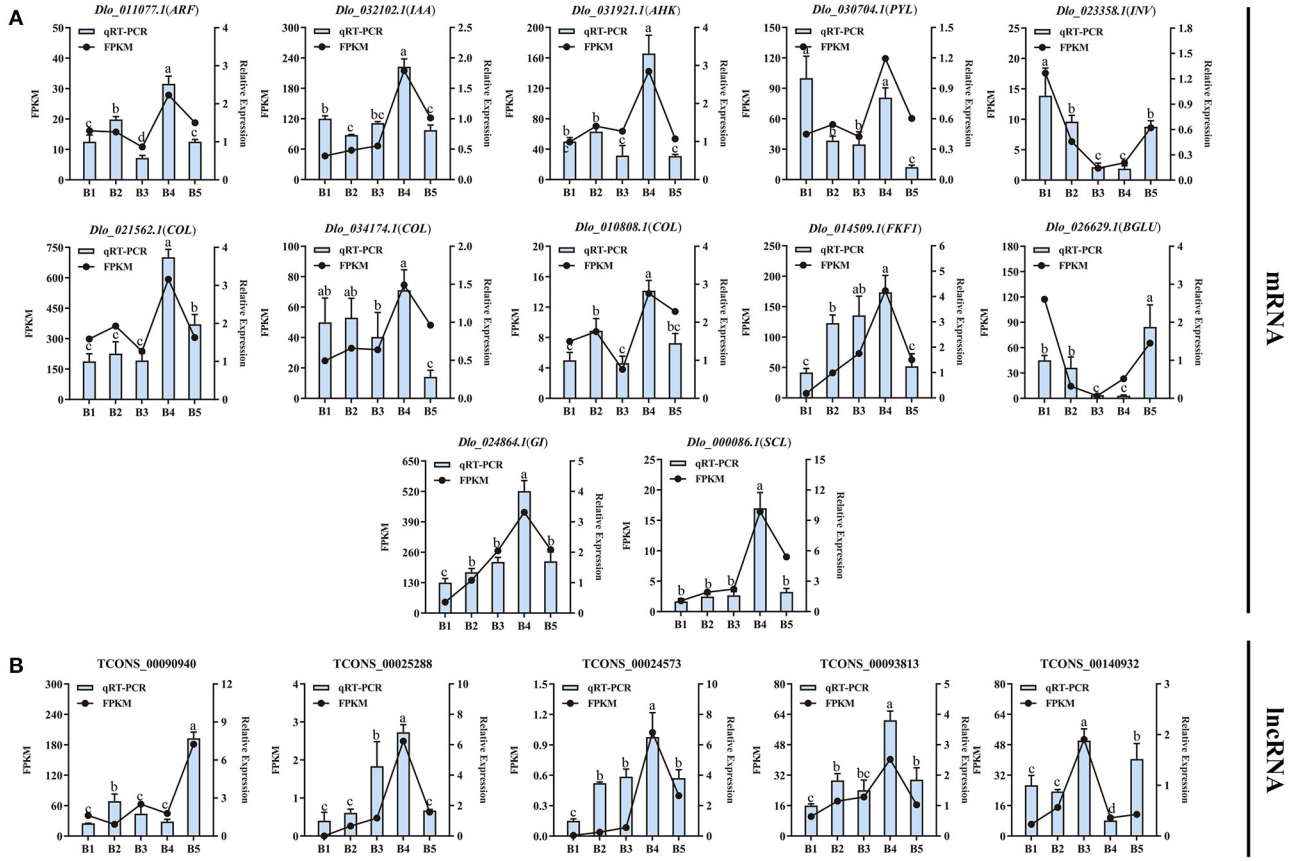


FIGURE 9 | Consistency analysis of RNA-Seq and qRT-PCR data. **(A)** Verification of the consistency of RNA-Seq data and qRT-PCR of 12 DEmRNAs. **(B)** Verification of the consistency of RNA-Seq data and qRT-PCR of 5 DElncRNAs. The different letters (a, b, c, d) show significant differences at the $p < 0.05$ level based on Duncan's test.

endogenous nutrition and hormones are important for the growth of fruiting branches. In addition, it is the fruiting season for "LDB" longan in September and October. Developing fruits provide a strong sink for photoassimilates. It is thought that depletion of photoassimilates, especially carbohydrates from the buds, may affect the number of floral buds (Goldschmidt et al., 1985; Goldschmidt, 1999; Shalom et al., 2014). Therefore, it is speculated that the sucrose accumulated at B1 stage may be partly supplied to the developing fruits. At B3–B4 stage, DEmRNAs were significantly enriched in all three pathways, especially genes in the circadian rhythm pathway were strongly up-regulated. November and December (B3 and B4 stages) were thought to be the period of physiology differentiation of floral bud induction (Yuan, 2016), when apical buds sense environmental changes, such as low temperature and short photoperiod, to ensure the stability of floral induction (Suttitanawat et al., 2012). Our results were consistent with the characteristics of physiological differentiation of floral induction, suggesting that apical buds respond to external environmental signals and initiate the internal physiological changes in this period. Meanwhile, our results also showed that sucrose and hormones were necessary for the physiological differentiation of floral buds. Previous studies have highlighted that sucrose-mediated signaling is interconnected with hormone-mediated floral induction, possibly as an additional part of flowering initiation (Wang et al., 2021).

Starch and Sucrose Metabolism-Related Genes Mediate Floral Induction in Longan

Carbohydrates are essential for plant growth and development, including flowering (Gibson, 2005; Matsoukas, 2014). B1 stage-related MEdarkred and MEblack modules significantly enriched a large number of sucrose metabolism-related genes (such as *SUS*, *INV*, and *HXK*), indicating that apical buds of longan mainly provided nutrients in the form of sucrose during B1 and B2 stages. Notably, four homologous genes of *HXK*, which act as sugar sensors sensing sugar levels and phosphorylation status and transmitting signals (Jang and Sheen, 1994; Rolland and Sheen, 2005), were only expressed at the B1 stage with spatiotemporal specificity, suggesting that *HXK* genes in longan participated in the sucrose-induced signal transduction at B1 stage and responded specifically to the growth and development of autumn shoots. Trehalose-6-phosphate (T6P) has been proposed as a proxy for carbohydrate status in plants (Wahl et al., 2013). T6P

level is positively correlated with sucrose level, and an increase in T6P level leads to an increase in sugar metabolism and signaling which initiates floral induction (Lastdrager et al., 2014). In our study, the expression levels of *TPS* genes, which were key genes involved in T6P biosynthesis pathway, were up-regulated during B1–B2 stage, suggesting that both sucrose and T6P act as the proxy for the carbohydrate status in fruiting branches of longan.

Starch is an important energy storage form that can be hydrolyzed into sugar for the utilization of floral buds (Li et al., 2013). The expression levels of key enzyme genes related to starch degradation and sucrose biosynthesis, such as *BAM*, *DPE,* and *SPS*, were up-regulated at B3 or B4 stage, leading to starch degradation and sucrose accumulation, which provided energy for the physiological differentiation of floral induction. A similar result was reported in apple (Xing et al., 2015). *BAM* genes (*AtBAM1*, *AtBAM3*) are required for meristem function in *Arabidopsis* (DeYoung et al., 2006). In the present study, *BAM* genes (*Dlo_026796.1*, *Dlo_029595.1*, *Dlo_003118.1*) had higher expression at B3 and B4 stages, suggesting that they may be involved in maintaining the normal development of longan floral bud meristem. Sugar levels and metabolism-related gene expression profiles revealed that sucrose is the initiation signal in apple floral induction, and sucrose accumulation has a positive influence on floral induction and development (Xing et al., 2015; Du et al., 2017). This suggests that sucrose plays an essential role in the regulation of bud growth and carbohydrate levels that induce sugar signal and promote floral induction in longan.

Circadian Clock and Photoperiod-Related Genes Play key Roles During Floral Induction in Longan

Plants possess a circadian clock that enables them to coordinate internal biological events with changes in external rhythm (Imaizumi, 2010). The photoperiod pathway mediates light and temporal information from the environment into the regulation of flowering time (Cheng and Wang, 2010). In perennial woody plants, the circadian clock may sense seasonal cues and induce flowering seasonally (Jia et al., 2014). In this study, *FKF1*, *GI*, *APRR5* (*Dlo_006275.1, Dlo_032811.1*), *COL*, and *SOC1* genes exhibited higher transcript accumulation levels at the B4 stage, indicating that circadian clock and photoperiod-mediated flowering induction is an important pathway in the flowering regulatory network in longan. *FKF1* and *GI* play important roles in circadian oscillation and flowering time regulation (Tootle et al., 2003; Mishra and Panigrahi, 2015). A previous report suggested that longan *GI* and *FKF1* were flowering promoters and may regulate the floral bud physiological differentiation by affecting endogenous auxin synthesis and transport (Huang et al., 2017). Consistent with that, in this study, *FKF1* was found to act as a central hub gene co-expressed with auxin signaling-related genes in the B4-related MEyellow module. In addition, two *CONSTANS-LIKE* (*COL*) family members (*Dlo_021562.1*, *Dlo_034174.1*) were identified in the photoperiod pathway, whose homolog (*AtCOL4*, *AtCOL9*) had an important role in regulating flowering time in *Arabidopsis* (Cheng and Wang, 2010; Steinbach, 2019). It has been reported that the deletion of a pair of *COL* genes may affect flowering time in lychee (Hu et al., 2022). *COL* is the central regulator in the photoperiod pathway, down-stream of the circadian control (Jang et al., 2008). In this study, longan two *COL* genes were found to be co-expressed with *FKF1* and *GI*, suggesting they may be under the direct regulation of *FKF1* and *GI* the same as in *Arabidopsis*. Moreover, *CO* was reported to be directly up-regulating the floral-pathways integrator *FT* gene (Valverde et al., 2004). However, the transcript of *FT* was not detected in our transcriptome database. Instead, *SOC1* was detected. *SOC1* is also a floral-pathways integrator, which is directly activated by *CO* in long photoperiod in *Arabidopsis* (Lee and Lee, 2010) and litchi (Zhang et al., 2014). These results suggest that *SOC1* may be the key flowering integrator in longan and *COL* (*Dlo_021562.1*, *Dlo_034174.1*) may control photoperiod floral induction by activating the expression of *SOC1*. *APRR5* coordinates and positively regulates flowering time by regulating the *CO*-dependent photoperiod pathway (Nakamichi et al., 2005, 2007). Two *APRR5* genes showed strong up-regulation at B4 stage, suggesting that they had a potential function in longan floral induction. Overall, key genes related to circadian clock and photoperiod pathway were identified in longan, circadian clock, and photoperiod regulatory pathway, which may play a key role in seasonal floral induction in longan.

Hormone Biosynthesis and Signaling Mediate Floral Induction in Longan

Phytohormones are signaling molecules induced by plant cells to receive specific environmental signals and regulate multiple aspects of plant physiological responses and development at low concentrations (Davis, 2009). Auxin is a key factor in inducing flower formation (Cho et al., 2017). The Medarkred, Meblack, MEyellow, and MEblue co-expression center networks all contained multiple auxin signaling-related genes, indicating that auxin plays an important role in different stages of longan floral induction. At the B4 stage, three auxin signaling-related genes (*ARF*, *Dlo_011077.1*, *Dlo_005415.2*; *AUX1*, *Dlo_006039.2*) in the co-expression center network of MEyellow modules showed increased expression. Moreover, three auxin signaling-related genes (*SAUR*, *Dlo_027349.1*, *Dlo_002885.1*; *GH3*, *Dlo_010164.3*) in the MEblue module showed decreased expression. Both *GH3* and *SAUR* genes have a negative feedback regulation of auxin signaling (Staswick, 2005; Nektarios et al., 2013). Several studies have indicated that high concentrations of IAA may have the effect of promoting the physiological differentiation of floral buds in longan (Su et al., 1997; Yang et al., 2019). Our results confirm that auxin may promote the floral bud differentiation in longan. The identified candidate gene was *ARF* (*Dlo_005415.2*), whose homolog (*AtARF2*) showed late flowering in *Arabidopsis* mutants (Okushima et al., 2005). As the core gene of auxin signal transduction pathway, *ARF* plays a synergistic role in multiple pathways mediated by auxin transduction signal (Peng et al., 2020). In addition, *ARFs* represent a point of cross-talk between ethylene and auxin signaling (Li et al., 2006). *ARF* (*Dlo_005415.2*, *Dlo_011077.1*) as hub genes were co-expressed with photoperiod-related genes (*COL*, *FKF1*), sucrose biosynthesis gene (*SPS*), and *ERF* genes (**Figure 4C**), indicating that *ARF* genes have potential

regulatory capacity of crosstalk in synergistic key pathways related to longan floral induction.

In the model plant *Arabidopsis*, GA plays an active role in flowering through a GA-dependent signaling pathway (Mutasa-Gottgens and Hedden, 2009). Some perennial woody fruit trees, including mango (Nakagawa et al., 2012) and apple (Wilkie et al., 2008; Zhang et al., 2019), have been shown to have a negative response for GAs in floral bud formation. Also, several studies have documented that the floral induction in longan requires low levels of GAs (Huang, 1996; Wang et al., 2015). As a key negative regulator of GA signaling, *DELLA* gene was down-regulated during B3 and B4 stages, which was inconsistent with previous results of the inhibition function of GAs in longan. Therefore, it may be a different regulatory mechanism of GAs with *Arobidopsis* during flowering initiation in longan, which needs more investigation.

Ethylene is widely distributed in plants and closely related to the process of plant flower formation (Johnson and Ecker, 1998). However, unlike auxin and GA, little attention has been paid to ETH function in floral induction of woody fruit trees. In contrast to other stages, DEmRNAs in ETH biosynthesis maintained low expression levels at the B4 stage. This finding showed that reduced ETH level at the B4 stage may favor the process of floral bud induction. *ERF* is the key transcription factor in ETH signaling pathway (Zhang et al., 2020a). In the central co-expression network of the MEblue module, *ERF* (*Dlo_015669.1*, *Dlo_007316.1*) genes were co-expressed with ETH biosynthesis-related genes, all of which showed down-regulated expression levels at B4 stage. Therefore, these two *ERFs* may function by responding to decreased ethylene levels and sense external environmental changes at B4 stage in longan.

Crosstalk Among Three Pathways

Longan floral induction requires coordination between external environmental conditions and various internal physiological factors, such as hormones and carbohydrates (Lin and Lu, 1992). It has been reported that exogenous auxin resistance was shown in *Arabidopsis hxk*/*gin2* mutants, while auxin-resistant mutants were insensitive to high glucose levels (Rolland and Sheen, 2005). In addition, the expression levels of auxin and trehalose biosynthesis-related genes in buds were simultaneously affected by fruit load in citrus (Shalom et al., 2014). These studies suggest a link between sugar and hormones. In our study, sucrose metabolism-related genes were co-expressed with hormone signaling-related genes in the MEdarkred and MEblack networks, revealing the crosstalk between these two pathways. An interaction of sucrose and hormones could take place at B1 and B2 stages, affecting the growth and development of autumn shoots of longan.

WGCNA co-expression network analysis revealed that three putative key pathways interacted at the B4 stage to form a regulatory network. Sugars are inducers of signals. Exogenous sucrose application can induce the expression of circadian clock-related genes, such as *AtGI* and *AtTOC1* (Dalchau et al., 2011). Endogenous sugars (such as sucrose and glucose) produced by photosynthesis can entrain circadian clock-related genes in *Arabidopsis* (Román et al., 2021). These findings suggest that sugars have crosstalk with circadian-related genes during flowering in plants. More evidence supporting the close association of flowering with sugar-related genes has been reported (Seo et al., 2011). The late-flowering phenotype of the *CO*-defective mutant is partly restored by sucrose feeding in darkness, suggesting that sucrose-mediated signals are incorporated into the photoperiod flowering pathway (Roldán et al., 1999). Consistent with that, in our study, rhythm genes in MEyellow and MEblue modules as central hubs were co-expressed with sucrose metabolism-related genes. Sucrose accumulation at the B4 stage may induce the upregulation of circadian rhythm-related genes. In addition, previous studies showed that the expression of a subset of genes involved in auxin biosynthesis, perception, and signaling was controlled by the clock (Covington and Harmer, 2007). *DlGI* and *DlFKF1* transgenic *Arabidopsis* showed that *DlGI* and *DlFKF1* may regulate auxin biosynthesis and transport (Huang et al., 2017). Consistent with this, *FKF1*, *GI*, and auxin-related genes (*ARF*, *Dlo_011077.1*, *Dlo_005415.2*) were co-expressed and showed high correlation at B4 stage, suggesting that circadian rhythm genes interact with the hormones signal transduction pathway and participate in longan seasonal floral induction.

LncRNAs Regulate Longan Floral Induction by Coordinating the Expression of Genes in Three key Pathways

Comparing the co-expression of neighboring and random lncRNA–mRNA pairs revealed that lncRNAs primarily regulate neighboring gene expression in a positive *cis*-regulatory manner. This *cis*-regulatory effect has also been observed in other plants, such as *Oryza sativa*, *Vitis vinifera*, and *Populus trichocarpa* (Zhang et al., 2020b; Huang et al., 2021; Ye et al., 2022). In contrast to the boundedness of neighboring *cis*-regulatory functions, the *trans*-regulatory functions of lncRNAs at the genome-scale theoretically have an infinite set of target genes (Ye et al., 2022). Therefore, studying lncRNA–mRNA pairs expressed in a similar pattern will provide more information. In this study, based on the fact that the number of co-transcribed lncRNA–mRNA pairs was much more than that of *cis*-regulated lncRNA–mRNA pairs, it is indicated that *trans*-regulation of lncRNA is the main way to regulate three key pathways.

Some lncRNAs have been reported to target significantly differentially expressed mRNAs with different functions as well as specifically target functional mRNAs, and lncRNAs may act as regulators (Xue et al., 2019). The same results were shown in our study. For instance, two lncRNAs (TCONS_00036970, TCONS_00063003) positively regulate the expression of both ethylene and sucrose metabolism-related genes (*BAM*, *Dlo_026796.1*, *Dlo_029595.1*; *ERF*, *Dlo_015669.1*; *ACS*, *Dlo_033533.1*; *BGLX, Dlo_024997*). DEmRNAs analysis above has shown the interaction among sucrose metabolism, rhythms, and hormones. These results provide more evidence that lncRNAs participate in the regulatory network of floral induction by positively regulating the expression of key players of three pathways. In addition, lncRNAs may bind to flowering TFs through binding motifs and thus affect

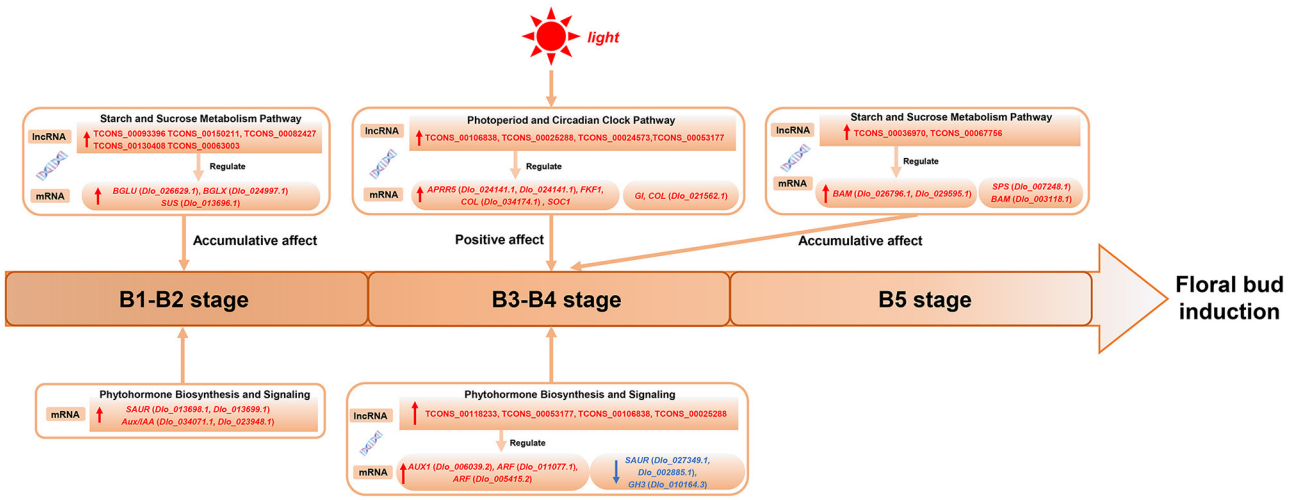


FIGURE 10 | A hypothetical model showing three genetic pathways regulating floral bud induction in longan.

floral bud differentiation. Similar results also appeared in the study of *Populus trichocarpa* (Ye et al., 2022). This specific mechanism of lncRNA–mRNA action needs to be further studied.

Hypothetical Model of Regulatory Pathways, key Genes and LncRNAs During Floral bud Induction in Longan

A hypothetical model for the regulation of longan floral induction *via* sugar, circadian, and hormone signaling and their crosstalk are shown in **Figure 10**; 21 hub-DEmRNAs and 12 hub-DElncRNAs involved in the three pathways are listed. In conclusion, at B1 stage the expression of sucrose metabolism-related genes (*BGLU*, *Dlo_026629.1*; *BGLX*, *Dlo_024997.1*; *SUS*) and auxin signaling negative feedback factors (*SAUR*, *Dlo_013698.1*, *Dlo_013699.1*; *Aux/IAA*, *Dlo_034071.1, Dlo_023948.1*) were up-regulated, which may be related to the growth and maturation of autumn shoots and fruit load. During physiological differentiation of floral buds, especially at the B4 stage, the circadian rhythm-related genes (*APRR5*, *FKF1*, *GI*, *COL*, *SOC1*) were co-expressed with sucrose metabolism-related genes (*SPS*, *BAM*) and auxin signal transduction-related genes (*ARF*, *Dlo_011077.1*, *Dlo_005415.2*; *AUX1*, *Dlo_006039.2*), actively participating in the regulation of the floral bud induction. High expression of auxin signaling genes may be associated with sucrose accumulation and positive response to circadian rhythm signals. Meanwhile, 12 hub-DElncRNAs in the three pathways were also showed, which affected gene expression in a positive regulatory manner and strengthened the connection of the three pathways.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: https://www.ncbi.nlm.nih.gov/, PRJNA830603.

AUTHOR CONTRIBUTIONS

LZ and SZ conceived the study and designed the experiments. FL, YZ, XW, SY, and TF conducted experiments and analyzed data. FL wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was financially supported by the Fujian Agriculture and Forestry University Science and Technology Innovation Fund (KFA20028A) and the Fujian Agriculture and Forestry University Outstanding Graduate Student Fund (1122YS010 and 1122YS01008).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.923183/full#supplementary-material

REFERENCES

Albani, M.C., Castaings, L., Wotzel, S., Mateos, J.L., Wunder, J., Wang, R., et al. (2012). PEP1 of Arabis alpina is encoded by two overlapping genes that contribute to natural genetic variation in perennial flowering. *PLoS Genet.* 8, e1003130. doi: 10.1371/journal.pgen.1003130

Bailey, T. L. (2021). STREME: accurate and versatile sequence motif discovery. *Bioinformatics* 18, 18. doi: 10.1093/bioinformatics/btab203

Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., and Dong, D. (2019). LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.* 47, D1034–D1037. doi: 10.1093/nar/gky905

Buske, A., and Bodén, B. (2010). Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics* 26, 860–866. doi: 10.1093/bioinformatics/btq049

Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202. doi: 10.1016/j.molp.2020.06.009

Cheng, X. F., and Wang, Z. Y. (2010). Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in Arabidopsis thaliana. *Plant J.* 43, 758–768. doi: 10.1111/j.1365-313X.2005.02491.x

Cho, L. H., Yoon, J., and An, G. (2017). The control of flowering time by environmental factors. *Plant J.* 90, 708–719. doi: 10.1111/tpj.13461

Covington, M. F., and Harmer, S. L. (2007). The circadian clock regulates auxin signaling and responses in arabidopsis. *Plos Biol.* 5, e222. doi: 10.1371/journal.pbio.0050222

Curaba, J., Singh, M. B., and Bhalla, P. L. (2014). miRNAs in the crosstalk between phytohormone signalling pathways. *J. Exp. Bot.* 65, 1425–1438. doi: 10.1093/jxb/eru002

Dalchau, N., Baek, S. J., Briggs, H. M., Robertson, F. C., Dodd, A. N., Gardner, M. J., et al. (2011). The circadian oscillator gene GIGANTEA mediates a long-term response of the Arabidopsis thaliana circadian clock to sucrose. *Proceed. Nat. Acad. Sci. United States Am.* 108, 5104–5109. doi: 10.1073/pnas.1015452108

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

de Montaigu, A., Toth, R., and Coupland, G. (2010). Plant development goes like clockwork. *Trends Genet.* 26, 296–306. doi: 10.1016/j.tig.2010.04.003

DeYoung, B. J., Bickle, K. L., Schrage, K. J., Muskett, P., Patel, K., and Clark, S. E. (2006). The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in Arabidopsis. *Plant J.* 45, 1–16. doi: 10.1111/j.1365-313X.2005.02592.x

Du, L., Qi, S., Ma, J., Xing, L., Fan, S., Zhang, S., et al. (2017). Identification of TPS family members in apple (Malus x domestica Borkh.) and the effect of sucrose sprays on TPS expression and floral induction. *Plant Physiol. Biochemistr.* 120, 10–23. doi: 10.1016/j.plaphy.2017.09.015

Fang, J., Zhang, F., Wang, H., Wang, W., Zhao, F., Li, Z., et al. (2019). Ef-cd locus shortens rice maturity duration without yield penalty. *Proc. Natl. Acad. Sci. U S A* 116, 18717–18722. doi: 10.1073/pnas.1815030116

Fu, Z. Y., Jia, T. Q., Peng, Y., Weheed, S., and Zeng, L. H. (2018). Cloning and function analysis of ELF4 homolog genes in dimocarpus longan. *Horticult. Plant J.* 45, 875–886. doi: 10.16420/j.issn.0513-353x.2017-0647

Gerttula, S., Zinkgraf, M., Muday, G. K., Lewis, D. R., Ibatullin, F. M., Brumer, H., et al. (2015). Transcriptional and hormonal regulation of gravitropism of woody stems in populus. *Plant Cell* 27, 2800–2813. doi: 10.1105/tpc.15.00531

Gibson, S. I. (2005). Control of plant development and gene expression by sugar signaling. *Curr. Opin. Plant Biol.* 8, 93–102. doi: 10.1016/j.pbi.2004.11.003

Goldschmidt, E. E. (1999). Carbohydrate supply as a critical factor for citrus fruit development and productivity. *Hortscience* 34, 1020–1024. doi: 10.21273/HORTSCI.34.6.1020

Goldschmidt, E. E., Aschkenazi, N., Herzano, Y., Schaffer, A. A., and Monselise, S. P. (1985). A role for carbohydrate levels in the control of flowering in citrus. *Scientia Horticulturae* 26, 159–166. doi: 10.1016/0304-4238(85)90008-1

Han, Y., Yang, H., and Jiao, Y. (2014). Regulation of inflorescence architecture by cytokinins. *Front. Plant Sci.* 5, 669. doi: 10.3389/fpls.2014.00669

Hao, Z., Fan, C., Cheng, T., Su, Y., Wei, Q., and Li, G. (2015). Genome-wide identification, characterization and evolutionary analysis of long intergenic noncoding RNAs in cucumber. *PLoS One* 10, e0121800. doi: 10.1371/journal.pone.0121800

Harrow, J., Frankish, A., Gonzalez, J. M., Tapanari, E., Diekhans, M., Kokocinski, F., et al. (2012). GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* 22, 1760–1774. doi: 10.1101/gr.135350.111

Heo, J. B., and Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331, 76–79. doi: 10.1126/science.1197349

Hu, G., Feng, J., Xiang, X., Wang, J., Salojarvi, J., Liu, C., et al. (2022). Two divergent haplotypes from a highly heterozygous lychee genome suggest independent domestication events for early and late-maturing cultivars. *Nat. Genet.* 54, 73–83. doi: 10.1038/s41588-021-00971-3

Hu, S., Wang, X., and Shan, G. (2016). Insertion of an Alu element in a lncRNA leads to primate-specific modulation of alternative splicing. *Nat. Struct. Mol. Biol.* 23, 1011–1019. doi: 10.1038/nsmb.3302

Huang, F. N., Fu, Z. Y., Zeng, L. H., and Morley-Bunker, M. (2017). Isolation and characterization of GI and FKF1 homologous genes in the subtropical fruit tree Dimocarpus longan. *Mol. Breed.* 37, 7. doi: 10.1007/s11032-017-0691-z

Huang, Q. W. (1996). Changes in endogenous hormone contents in relation to flower bud differentiation and on-year or off-year fruiting of longan. *J. Tropic. Subtropic. Bot.* 2, 58–62.

Huang, X. P., Zhang, H. Y., Wang, Q., Guo, R., Wei, L. X., Song, H. Y., et al. (2021). Genome-wide identification and characterization of long non-coding RNAs involved in flag leaf senescence of rice. *Plant Mol. Biol.* 105, 655–684. doi: 10.1007/s11103-021-01121-3

Imaizumi, T. (2010). Arabidopsis circadian clock and photoperiodism: time to think about location. *Curr. Opin. Plant Biol.* 13, 83–89. doi: 10.1016/j.pbi.2009.09.007

Jang, J., and Sheen, J. (1994). Sugar sensing in higher plants. *Plant Cell* 6, 1665–1679. doi: 10.1105/tpc.6.11.1665

Jang, S., Marchal, V., Panigrahi, K. C., Wenkel, S., Soppe, W., Deng, X. W., et al. (2008). Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. *EMBO J.* 27, 1277–1288. doi: 10.1038/emboj.2008.68

Jia, T., Wei, D., Meng, S., Allan, A. C., and Zeng, L. (2014). Identification of regulatory genes implicated in continuous flowering of longan (Dimocarpus longan L.). *PLoS ONE* 9, e114568. doi: 10.1371/journal.pone.0114568

Jiang, H., Ma, R., Zou, S., Wang, Y., Li, Z., and Li, W. (2017). Reconstruction and analysis of the lncRNA-miRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in rheumatoid arthritis. *Mol. Biosyst.* 13, 1182–1192. doi: 10.1039/C7MB00094D

Jin, J. J., Liu, J., Wang, H., Wong, L., and Chua, N. H. (2013). PLncDB: plant long non-coding RNA database. *Bioinformatics* 29, 1068–1071. doi: 10.1093/bioinformatics/btt107

Johnson, P. R., and Ecker, J. R. (1998). The ethylene gas signal transduction pathway: a molecular perspective. *Ann. Rev. Genet.* 32, 227–254. doi: 10.1146/annurev.genet.32.1.227

Jue, D., Sang, X., Liu, L., Shu, B., Wang, Y., Xie, J., et al. (2018). The ubiquitin-conjugating enzyme gene family in longan (dimocarpus longan lour.): genome-wide identification and gene expression during flower induction and abiotic stress responses. *Molecules* 23, 3. doi: 10.3390/molecules23030662

Ke, L. L., Zhou, Z. W., Xu, X. W., Wang, X., Liu, Y. L., Xu, Y. T., et al. (2019). Evolutionary dynamics of lincRNA transcription in nine citrus species. *Plant J.* 98, 912–927. doi: 10.1111/tpj.14279

Kim, J. J., Lee, J. H., Kim, W., Jung, H. S., Huijser, P., and Ahn, J. H. (2012). The microRNA156-squamosa promoter binding protein-like3 module regulates ambient temperature-responsive flowering via flowering locus T in Arabidopsis. *Plant Physiol.* 159, 461–478. doi: 10.1104/pp.111.192369

Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172, 393–407. doi: 10.1016/j.cell.2018.01.011

Kurokura, T., Mimida, N., Battey, N. H., and Hytonen, T. (2013). The regulation of seasonal flowering in the Rosaceae. *J. Exp. Bot.* 64, 4131–4141. doi: 10.1093/jxb/ert233

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9, 559. doi: 10.1186/1471-2105-9-559

Lastdrager, J., Hanson, J., and Smeekens, S. (2014). Sugar signals and the control of plant growth and development. *J. Exp. Bot.* 65, 799–807. doi: 10.1093/jxb/ert474

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J Exp. Bot.* 61, 2247–2254. doi: 10.1093/jxb/erq098

Li, J., Dai, X., and Zhao, Y. (2006). A role for auxin response factor 19 in auxin and ethylene signaling in arabidopsis. *Plant Physiol.* 140, 899–908. doi: 10.1104/pp.105.070987

Li, S. G., Hong, J. W., Zhang, L., and Yang, Z. Q. (2013). Carbon nutrition differences between on-season and off-season longan in flowering *Guangdong Agricult. Sci.* 40, 43-45+50. doi: 10.16768/j.issn.1004-874x.2013.24.033

Lin, S., and Lu, M. Y. (1992). Formation and cultivation of fruiting parent branches of longan. *Guangxi Agricult. Sci.* 05, 209–212.

Link, H. (2000). Significance of flower and fruit thinning on fruit quality. *Plant Growth Regulat.* 31, 17–26. doi: 10.1023/A:1006334110068

Liu, X., Hao, L., Li, D., Zhu, L., and Hu, S. (2015). Long non-coding RNAs and their biological roles in plants. *Genomics Proteom. Bioinform.* 13, 137–147. doi: 10.1016/j.gpb.2015.02.003

Long, Y., Wang, X., Youmans, D. T., and Cech, T. R. (2017). How do lncRNAs regulate transcription? *Sci. Adv.* 3, eaao2110. doi: 10.1126/sciadv.aao2110

Lyons, R., Rusu, A., Stiller, J., Powell, J., Manners, J. M., and Kazan, K. (2015). Investigating the Association between Flowering Time and Defense in the Arabidopsis thaliana-Fusarium oxysporum Interaction. *PLoS One* 10, e0127699. doi: 10.1371/journal.pone.0127699

Matsoukas, I. G. (2014). Interplay between sugar and hormone signaling pathways modulate floral signal transduction. *Front. Genet.* 5, 218. doi: 10.3389/fgene.2014.00218

Mishra, P., and Panigrahi, K. C. (2015). GIGANTEA - an emerging story. *Front. Plant. Sci.* 6, 8. doi: 10.3389/fpls.2015.00008

Mutasa-Gottgens, E., and Hedden, P. (2009). Gibberellin as a factor in floral regulatory networks. *J. Exp. Bot.* 60, 1979–1989. doi: 10.1093/jxb/erp040

Nakagawa, M., Honsho, C., Kanzaki, S., Shimizu, K., and Utsunomiya, N. (2012). Isolation and expression analysis of FLOWERING LOCUS T-like and gibberellin metabolism genes in biennial-bearing mango trees. *Scientia Horticulturae* 139, 108–117. doi: 10.1016/j.scienta.2012.03.005

Nakamichi, N., Kita, M., Ito, S., Yamashino, T., and Mizuno, T. (2005). PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, together play essential roles close to the circadian clock of Arabidopsis thaliana. *Plant Cell Physiol.* 46, 686–698. doi: 10.1093/pcp/pci086

Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., et al. (2007). Arabidopsis Clock-Associated Pseudo-Response Regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical constans-dependent photoperiodic pathway. *Plant Cell Physiol.* 48, 822. doi: 10.1093/pcp/pcm056

Nektarios, M. M., Karolina, B., Bram, V. L., Kumud, S., Susanna, C., Jean-Pierre, V., et al. (2013). Characterization of a Small Auxin-Up RNA (SAUR)-like gene involved in arabidopsis thaliana development. *Plos One* 8, e82596. doi: 10.1371/journal.pone.0082596

Okushima, Y., Mitina, I., Quach, H. L., and Theologis, A. (2005). AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. *Plant J.* 43, 29–46. doi: 10.1111/j.1365-313X.2005.02426.x

Paytuvi Gallart, A., Hermoso Pulido, A., Anzar Martinez de Lagran, I., Sanseverino, W., and Aiese Cigliano, R. (2016). GREENC: a Wiki-based database of plant lncRNAs. *Nucleic Acids Res.* 44, D1161–1166. doi: 10.1093/nar/gkv1215

Peng, Y., Fang, T., Zhang, Y. Y., Zhang, M. Y., and Zeng, L. H. (2020). Genome-wide identification and expression analysis of auxin response factor (arf) gene family in longan (Dimocarpus longan L.). *Plants (Basel)* 9, 221. doi: 10.3390/plants9020221

Pin, P. A., and Nilsson, T. (2012). The multifaceted roles of FLOWERING LOCUS T in plant development. *Plant Cell Environ.* 35, 1742–1755. doi: 10.1111/j.1365-3040.2012.02558.x

Povey, S., Lovering, R., Bruford, E., Wright, M., Lush, M., and Wain, H. (2001). The HUGO gene nomenclature committee (HGNC). *Hum. Genet.* 109, 678–680. doi: 10.1007/s00439-001-0615-0

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

Roldán, M., Gómez-Mena, C., Ruiz-García, L., Salinas, J., and Martínez-Zapater, J. (1999). Sucrose availability on the aerial part of the plant promotes morphogenesis and flowering of Arabidopsis in the dark. *Plant J.* 20, 581–590. doi: 10.1046/j.1365-313X.1999.00632.x

Rolland, F., and Sheen, J. (2005). Sugar sensing and signalling networks in plants. *Biochem. Soc. Trans.* 33, 269–271.

Román, N., Li, X., Deng, D., Davey, J. W., and Haydon, M. J. (2021). Superoxide is promoted by sucrose and affects amplitude of circadian rhythms in the evening. *Proceed. Nat. Acad. Sci.* 118, e2020646118. doi: 10.1073/pnas.2020646118

Seo, P. J., Ryu, J., Kang, S. K., and Park, C. M. (2011). Modulation of sugar metabolism by an INDETERMINATE DOMAIN transcription factor contributes to photoperiodic flowering in Arabidopsis. *Plant J.* 65, 418–429. doi: 10.1111/j.1365-313X.2010.04432.x

Serrano, G., Herrera-Palau, R., Romero, J. M., Serrano, A., Coupland, G., and Valverde, F. (2009). Chlamydomonas CONSTANS and the evolution of plant photoperiodic signaling. *Curr. Biol.* 19, 359–368. doi: 10.1016/j.cub.2009.01.044

Shalom, L., Samuels, S., Zur, N., Shlizerman, L., Doron-Faigenboim, A., Blumwald, E., et al. (2014). Fruit load induces changes in global gene expression and in abscisic acid (ABA) and indole acetic acid (IAA) homeostasis in citrus buds. *J. Exp. Bot.* 65, 3029–3044. doi: 10.1093/jxb/eru148

Shimakawa, A., Shiraya, T., Ishizuka, Y., Wada, K. C., Mitsui, T., and Takeno, K. (2012). Salicylic acid is involved in the regulation of starvation stress-induced flowering in Lemna paucicostata. *J. Plant Physiol.* 169, 987–991. doi: 10.1016/j.jplph.2012.02.009

Staswick, P. (2005). Characterization of an Arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. *Plant Cell* 17, 616–627. doi: 10.1105/tpc.104.026690

Steinbach, Y. (2019). The Arabidopsis thaliana CONSTANS-LIKE 4 (COL4)–a modulator of flowering time. *Front. Plant Sci.* 10, 651. doi: 10.3389/fpls.2019.00651

Su, M. H., Liu, Z. C., and Zhuang, Y. M. (1997). Effect of endogenous hormone content on flower bud differentiation in Shuichang Longan Fruit Shoots. *J. Tropic. Crops* 02, 66–71.

Sun, Q., Csorba, T., Skourti-Stathaki, K., Proudfoot, N. J., and Dean, C. (2013). R-loop stabilization represses antisense transcription at the Arabidopsis FLC locus. *Science* 340, 619–621. doi: 10.1126/science.1234848

Suttitanawat, P., Sruamsiri, P., and Sringarm, K. (2012). Changes in cytokinins concentrations during induction period of longan cv. Daw in sand culture. *Int. J. Agricult. technol.* 8, 2353–2362.

Tootle, T. L., Silver, S. J., Davies, E. L., Newman, V., Latek, R. R., Mills, I. A., et al. (2003). The transcription factor Eyes absent is a protein tyrosine phosphatase. *Nature* 426, 299–302. doi: 10.1038/nature02097

Turnbull, C. (2011). Long-distance regulation of flowering time. *J. Exp. Bot.* 62, 4399–4413. doi: 10.1093/jxb/err191

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003–1006. doi: 10.1126/science.1091761

Wahl, V., Ponnu, J., Schlereth, A., Arrivault, S., Langenecker, T., Franke, A., et al. (2013). Regulation of flowering by trehalose-6-phosphate signaling in Arabidopsis thaliana. *Science* 339, 704–707. doi: 10.1126/science.1230406

Wang, C. C., and Ke, G. W. (1992). Study on flower bud morphological differentiation of Longan. *J. Fujian Acad. Agricult. Sci.* 01, 55–58.

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, M., Le Gourrierec, J., Jiao, F., Demotes-Mainard, S., Perez-Garcia, M. D., Oge, L., et al. (2021). Convergence and divergence of sugar and cytokinin signaling in plant development. *Int. J. Mol. Sci* 22, 3. doi: 10.3390/ijms22031282

Wang, W., Su, M. H., WU, S. H., and Chang, Q. (2015). Investigation of endogenous growth substances level changes in off season flower induction of longan (dimocarpus longan lour.) Trees. *Chinese J. Tropic. Crops* 36, 252–257. doi: 10.3389/fpls.2021.670587

Wei, S., Chen, H., Dzakah, E. E., Yu, B., Wang, X., Fu, T., et al. (2019). Systematic evaluation of C. *elegans* lincRNAs with CRISPR knockout mutants. *Genom. Biol.* 20, 7. doi: 10.1186/s13059-018-1619-6

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

Xing, L. B., Zhang, D., Li, Y. M., Shen, Y. W., Zhao, C. P., Ma, J. J., et al. (2015). Transcription profiles reveal sugar and hormone signaling pathways mediating flower induction in apple (malus domestica borkh.). *Plant Cell Physiol.* 56, 2052–2068. doi: 10.1093/pcp/pcv124

Xue, L., Sun, M., Wu, Z., Yu, L., and Jiang, F. (2019). LncRNA regulates tomato fruit cracking by coordinating gene expression via a hormone-redox-cell wall network. *BMC Plant Biol.* 20, 162. doi: 10.1186/s12870-020-02373-9

Yamaguchi, A., and Abe, M. (2012). Regulation of reproductive development by non-coding RNA in Arabidopsis: to flower or not to flower. *J. Plant Res.* 125, 693–704. doi: 10.1007/s10265-012-0513-7

Yamaguchi, N., Winter, C. M., Wu, M. F., Kanno, Y., and Wagner, D. (2014). Gibberellin Acts positively then negatively to control onset of flower formation in arabidopsis. *Science* 344, 638–641. doi: 10.1126/science.1250498

Yang, Z. Q., Huang, X. M., Chen, Y. Y., Hong, J. W., Li, S. G., Luo, J. B., et al. (2019). Changes of endogenous hormone level during flower development in on-season and off-season longan. *J.Anhui Agric.Sci.* 47, 47–74. doi: 10.3969/j.issn.0517-6611.2019.09.015

Ye, X., Wang, S., Zhao, X., Gao, N., Wang, Y., Yang, Y., et al. (2022). Role of lncRNAs in cis- and trans-regulatory responses to salt in Populus trichocarpa. *The Plant J.* 22, 15714. doi: 10.1111/tpj.15714

You, X. R., Wang, L. X., Liang, W. Y., Gai, Y. H., Wang, X. Y., and Chen, W. (2012). Floral reversion mechanism in longan (Dimocarpus longan Lour.) revealed by proteomic and anatomic analyses. *J Proteomics* 75, 1099–1118. doi: 10.1016/j.jprot.2011.10.023

Yuan, J. P., Zhang, Y., Dong, J. S., Sun, Y. Z., Lim, B. L., Liu, D., et al. (2016). Systematic characterization of novel lncRNAs responding to phosphate starvation in Arabidopsis thaliana. *BMC Genomics* 17, 655. doi: 10.1186/s12864-016-2929-2

Yuan, T. (2016). Growth and development characteristics and cultivation techniques of "Lidongben" of extra-late Maturing Longan. *South China Fruits* 45, 138–140. doi: 10.13938/j.issn.1007-1431.20150537

Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, 17. doi: 10.2202/1544-6115.1128

Zhang, H., Pan, X., Liu, S., Lin, W., and Zhang, X. (2020a). Genome-wide analysis of AP2/ERF transcription factors in pineapple reveals functional divergence during flowering induction mediated by ethylene and floral organ development. *Genomics* 113, 2. doi: 10.1016/j.ygeno.2020.10.040

Zhang, H. L., Wang, Z. G., Yu, Y. H., and Guo, D. L. (2020b). Genome-wide identification and characterization of long non-coding RNAs involved in grape berry ripening. *J. Berry Res.* 10, 1–22. doi: 10.3233/JBR-190518

Zhang, H. N., Wei, Y. Z., Shen, J. Y., Lai, B., Huang, X. M., Ding, F., et al. (2014). Transcriptomic analysis of floral initiation in litchi (Litchi chinensisSonn.) based on de novo RNA sequencing. *Plant Cell Rep.* 33, 1723–1735. doi: 10.1007/s00299-014-1650-3

Zhang, S., Gottschalk, C., and van Nocker, S. (2019). Genetic mechanisms in the repression of flowering by gibberellins in apple (Malus x domestica Borkh.). *BMC Genom.* 20, 747. doi: 10.1186/s12864-019-6090-6

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 16 June 2022
doi: 10.3389/fpls.2022.929892

The VvWRKY37 Regulates Bud Break in Grape Vine Through ABA-Mediated Signaling Pathways

Feng-Pan Wang[1,2†], Pan-Pan Zhao[3†], Lei Zhang[4], Heng Zhai[2], Muhammad Abid[1] and Yuan-Peng Du[2]*

[1] Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, China, [2] National Key Laboratory of Crop Biology, National Research Center for Apple Engineering and Technology, College of Horticulture Science and Engineering, Shandong Agricultural University, Tai'an, China, [3] Guangdong Provincial Key Laboratory for Plant Epigenetics, College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China, [4] College of Biological and Enology Engineering, Taishan University, Tai'an, China

Dormancy is a common survival strategy in plants to temporarily suspend visible growth under unsuitable conditions. The elaborate mechanism underlying bud break in perennial woody plants is gradually illustrated. Here, we identified a grape vine WRKY transcription factor, *VvWRKY37*, which was highly expressed in dormant buds. It was particularly induced by the application of exogenous abscisic acid, and depressed on exposure to gibberellin and low temperature (4°C) stress at the transcript level. The yeast one-hybrid assay confirmed that VvWRKY37 had a transcriptional activity. Ectopic over-expression of *VvWRKY37* significantly delayed bud break of transgenic poplar plants. As an ABA-inducible gene, *VvWRKY37* also depressed the expression of ABA catabolic gene *CYP707A*s and enhanced the accumulation of endogenous ABA in transgenic poplar plants. The molecular pieces of evidence showed that VvWRKY37 preferentially recognized and bound W-box 5′-G/CATTGACT/C/G-3′ *cis*-element *in vitro*. Additionally, VvABI5 and VvABF2 acted as the upstream transcriptional activators of *VvWRKY37 via* protein-DNA interactions. Taken together, our findings provided valuable insights into a new regulatory mechanism of WRKY TF by which it modulates bud break through ABA-mediated signaling pathways.

Keywords: grapevine, bud break, WRKY, abscisic acid, transcriptional regulation

OPEN ACCESS

Edited by:
Shunquan Lin,
South China Agricultural University, China

Reviewed by:
Walid El Kayal,
American University of Beirut, Lebanon
Xiping Wang,
Northwest A&F University, China
Jianzhao Li,
Ludong University, China

****Correspondence:***
Yuan-Peng Du
duyuanpeng001@163.com

† These authors have contributed equally to this work

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

Received: *27 April 2022*
Accepted: *24 May 2022*
Published: *16 June 2022*

Citation:
Wang F-P, Zhao P-P, Zhang L, Zhai H, Abid M and Du Y-P (2022) The VvWRKY37 Regulates Bud Break in Grape Vine Through ABA-Mediated Signaling Pathways. Front. Plant Sci. 13:929892. doi: 10.3389/fpls.2022.929892

INTRODUCTION

The growth and development of sessile plants is typically affected by climatic changes. Consequently, plants have evolved regulatory mechanisms to thrive well under unsuitable environmental conditions, i.e., freezing temperature. Perennial plants have developed a survival strategy that enables them to suspend and resume growth activities under cyclic changes in the environment. Meristem dormancy is a common strategy in most temperate perennial plants for their survival, development, and morphogenesis. Vegetative or floral buds of perennial plants bear pivotal responsibilities for growth and reproduction after dormancy release and thus ensure the sustainability of the plants after enduring unfavorable growth conditions. After dormancy establishment, endodormancy requires sufficient chilling intensity or other stimuli treatment (such as hydrogen cyanamide and sodium azide) before transitioning to ecodormancy (Xin and Browse, 2010). Ecodormancy can restore growth after a shift to growth-inductive conditions, indicating bud break. The signals originating from developmental (i.e., phytohormones), physiological (i.e.,

water, sugar, and phytochrome), and environmental (i.e., day length and temperature) factors play essential roles in the complex crosstalk regulating the bud dormancy and bud break (Chmielewski et al., 2017; Singh et al., 2017; Liu and Sherif, 2019).

So far, bud dormancy at the physiological level is well-studied (Parada et al., 2016; Chmielewski et al., 2017; Singh et al., 2017; Rubio et al., 2019), but the molecular and genetic mechanisms of the signaling networks that regulate dormancy and bud break are obscure. Previously, researchers have made numerous attempts to dissect genetic regulatory mechanisms underlying bud dormancy and break (Yordanov et al., 2014; Busov et al., 2016; Singh et al., 2018; Tylewicz et al., 2018; Artlip et al., 2019; Yang et al., 2019). Short-day length and low temperature synergistically promoted bud endodormancy and abscisic acid (ABA) accumulation in plants (Singh et al., 2017; Liu and Sherif, 2019). The ABA content gradually increased during bud dormancy establishment and decreased during endodormancy (Chmielewski et al., 2017). The overaccumulation of ABA prevented precocious growth during bud dormancy by blocking growth-related intracellular signaling cascades (Tylewicz et al., 2018). The ABA deficient mutant plants exhibited impaired dormancy potential (Singh et al., 2018), whereas ABA sufficient mutant and transgenic lines showed enhanced dormancy phenotypes (Yang et al., 2020). The ABA levels in winter buds increased in autumn and reached to a maximum level during endodormancy, and then persistently declined after endodormancy in grape plants (Rubio et al., 2019). The grape vines overexpressing the ABA catabolic gene *VvA8H-CYP707A4* gene enhanced ABA catabolism and bud break (Zheng et al., 2018a). All in all, an association between ABA content and bud behavior strengthens the argument for the regulatory role of ABA in bud dormancy induction, duration, and release.

Several studies have revealed the regulatory role of dormancy-related transcriptional factors (TFs), including dormancy-associated MADS-box (DAM) in aspen (Singh et al., 2018), pear (Yang et al., 2020), kiwifruit (Wu et al., 2017); APETALA2/Ethylene responsive factor in poplar (Yordanov et al., 2014), grape, and apple (Busov et al., 2016), Chinese cherry (Zhu et al., 2021); C-repeat binding factor (CBF/DREB) in peach (Artlip et al., 2019), pear (Li et al., 2019), bZIP-like TFs (ABFs) in pear (Yang et al., 2020), grape (Zheng et al., 2015); and WRKY TFs in peach (Chen et al., 2016). Interestingly, the above-mentioned dormancy-responsive TFs are directly or indirectly involved in dormancy-related phytohormone signal transduction cascades (Liu and Sherif, 2019). The expression patterns of WRKY TFs were correlated with bud dormancy in fruit trees (Chen et al., 2016; Vimont et al., 2019). Additionally, previous significant advances in ABA signaling studies have suggested that WRKY TF plays an essential role in ABA-responsive signaling networks (Rushton et al., 2012). In *Arabidopsis*, WRKY33 impeded ABA accumulation *via* transcriptional repression of the ABA biosynthesis genes *NCED3* and *NCED5* and activating the ABA catabolic gene *CYP707A3* in pathogen-induced defense response (Liu et al., 2015). It is reported that WRKY41 directly targets *ABI3* to trigger ABA signaling, which is beneficial for seed dormancy (Ding et al., 2014). Besides, WRKY6 positively regulated ABA-mediated seed germination by directly downregulating *RAV1* gene expression (Huang et al., 2016). OsWRKY29 represses seed dormancy in rice caused by the inhibitory effect of ABA signaling (Zhou et al., 2020).

Here, a grape WRKY TF (*VvWRKY37)* was characterized for its regulatory role in dormant winter bud. Moreover, we showed that the transcript abundance of *VvWRKY37* was induced by exogenous application of ABA and repressed on exposure to exogenous application of gibberellin (GA) and cold temperature. It is widely accepted that phytohormone ABA and cold temperatures are associated with plant dormancy in perennial plants. The present study provided evidence that VvWRKY37 is involved in bud dormancy and break *via* an ABA-mediated signaling pathway. This study also uncovered a VvWRKY37-mediated regulatory network by which GA and/or low-temperature exposure probably played respective roles in regulating bud dormancy and break.

MATERIALS AND METHODS

Plant Materials and Treatments

The self-rooted grape cultivar "Cabernet Sauvignon" (*Vitis vinifera*) from a vineyard located in Tai'an city, Shandong province, China (N36°09′49.24″, E117°08′53.73″) was used in this study. All berries were sampled from various 5-year-old grape vines cultivated using a vertical trellis system at the following stages: 7 days post-anthesis (dpa), (pepper-corn size berry), 35 dpa (pre-veraison berry), 70 dpa (veraison berry), and 105 dpa (harvest-ripe berry). Each berry was immediately peeled and deseeded. The skin, pulp, and seeds were frozen in liquid nitrogen and stored at −80°C except for the fruit at 7 dpa because of their small berries, making it difficult to separate the various tissues. The bud dormancy phenotyping was done for various dormancy-related parameters under environmental conditions. The dormancy depth was measured by using a previously described method (Pérez et al., 2007). The time period for the phase transition of endodormancy to ecodormancy was determined from December 4, 2015 to January 5, 2016 (**Supplementary Figure 1**). The appearance of visible green tissue at the tip of the winter bud was indicative of bud break. The data for bud break was collected on April 2, 2016 in spring. The winter buds were collected at the following phases: leaf fall in autumn (October 15, 2015), endodormancy (November 24, 2015), phase transition from endodormancy to ecodormancy (December 24, 2015), ecodormancy (February 15, 2016), and bud break in spring (April 2, 2016). The scales were quickly removed from the winter buds before freezing in liquid nitrogen.

Detached canes with more than 10 buds were sampled from a vineyard at the beginning of leaf fall. The experiment was repeated three times, and each replication contains 10 canes. The bottoms of canes were soaked in the following solution: ABA (10 μM), GA_3 (5 μM), and mock solution (ethanol, 0.06%), respectively. The canes were placed in a growth chamber set at 25°C/20°C temperature and 16 h/8 h

(light/dark) photoperiod. For low-temperature treatment, the bottoms of the canes were soaked in sterile water. The scales were carefully removed from each bud, and then the canes were transferred to low-temperature conditions (4°C) or normal conditions (20°C) as control. More than 20 buds were sampled for each time point. The scales were quickly removed from the winter buds before freezing in liquid nitrogen. Total RNA was extracted for analyzing transcript abundance analysis.

The poplar cultivar "84K" (*Populus alba* × *P. glandulosa*) was used as a wild genotype. *In vitro*-grown poplar shoots were sub-cultured every month using a maintenance medium, and the normal MS medium (20 g/L sucrose, 7. g/L agar powder, pH = 5.8) was supplemented with phytohormone NAA at 0.1 mg/L. The newly cultured shoots were grown at 25°C/20°C under a 16 h/8 h (light/dark) photoperiod in transparent glass bottles.

Plasmid Constructions and Plant Transformations

Standard molecular biological techniques and homologous recombination technology (Vazyme, Nanjing, China) were employed for gene cloning and plasmid construction. The coding sequence of *VvWRKY37* without stop codon was cloned from the first-strand cDNA, generating an in-frame C-terminal fusion to the GFP gene downstream of CaMV 35S promoter (*35S:VvWRKY37-GFP*). The recombinant plasmid was introduced into the LBA4404 strain of *Agrobacterium tumefaciens*.

The 1,165 bp upstream promoter fragment of *VvWRKY37* was amplified using grape vine (Cabernet Sauvignon) genomic DNA. The DNA fragment was inserted into the pCAMBIA1300-GN vector to generate the *pVvWRKY37*:*GUS* construct. Finally, the construct was introduced into the GV3101 strain of *Agrobacterium tumefaciens* and transformed into wild-type *A. thaliana* plants *via* the agrobacterium-mediated floral dip method. The T_0 seedlings were surface sterilized and screened on selective media. The well-grown resistant seedlings were transplanted and identified *via* genotyping using gene-specific primer pairs. The T_3 homozygous transgenic lines were used for subsequent experiments.

Leaves from *in vitro*-grown poplar "84K" shoot were used for transgenic plant regeneration. Fully expended young leaves were collected and co-cultivated with *Agrobacterium tumefaciens* strain LBA4404 harboring *35S:VvWRKY37-GFP* vector. The leaves were kept under dark conditions for 2 days and transferred into a regeneration medium containing 50 mg/L kanamycin and 250 mg/L cefotaxime. The regenerated shoots with a minimum of two expended leaves were cut and transferred into a rooting medium (kanamycin, 50 mg/L; cefotaxime, 250 mg/L; and NAA, 0.1 mg/L). Well-rooted plantlets were transplanted in plastic pots (10.5-cm height × 10.5-cm diameter) filled with a 50% peat and 50% perlite mixture. The pots were then covered with plastic bags for 1 week and incubated under constant growth conditions [25°C/20°C and a 16-h/8-h (light/dark) photoperiod].

Total RNA Extraction, RT-qPCR, and Semi-Quantitative RT-PCR

Plant total RNA was extracted using the CTAB solution [2% cetyltrimethyl ammonium bromide, 2.5% polyvinyl pyrrolidone K30, 2-M NaCl, 100-mM Tris–HCl, 25-mM EDTA-Na_2, 0.05% spermidine, 2% β-mercaptoethanol (added just before use)] as described previously with slight modifications (Gambino et al., 2008), or TRIzol™ Reagent (Thermo Fisher, Waltham, United States) by following the manufacturer's instructions. Microspectrophotometer NanoDrop One (Thermo Fisher, Waltham, United States) was used to quantify the total RNA, and the RNA integrity was confirmed using 1% agarose gel electrophoresis. The first-strand cDNA was synthesized from 1 μg of total RNA by PrimeScript™ RT Reagent Kit with genome DNA Eraser by following the manufacturer's protocol (TaKaRa, Tokyo, Japan). The stock solution of the first-stand cDNA was diluted 20 to 40 times with sterilized ddH_2O for RT-qPCR analysis. Then, RT-qPCR was carried out in a 96-well plate using Bio-Rad CFX96 (Bio-rad, Hercules, United States). Each well contained a total volume of 14-μl reaction mixtures consisting of 1 μl of cDNA, 7 μl of *TransStart*® Top Green qPCR SuperMix (TransGen, Beijing, China), 0.5 μl of each primer (10 mM), and 5 μl of sterilized ddH_2O. The PCR amplification procedure was as follows: initial denaturation at 94°C for 30 s, followed by 40 cycles at 94°C for 5 s, and 60°C for 30 s. Dissociation melting curve analysis was used to verify the PCR product specificity. Three housekeeping genes, *Vvactin*, *AtGAPDH,* and *PtUbi*, were used as internal references for grape, *Arabidopsis,* and poplar, respectively. The expression levels for genes were calculated by using the $2^{-\Delta\Delta Ct}$ method. The gene-specific primer pairs were synthesized using GENEWIZ (Suzhou, China), and the sequences for primer pairs are given in **Supplementary Table 1**.

Yeast One-Hybrid Assay

A previously described method was used to investigate the transcriptional activity of VvWRKY37 and its binding activities to downstream target genes (Wang F. et al., 2019). The full-length coding sequence of *VvWRKY37* was inserted into the pGADT7 vector to generate the effect or construct, and binding sequence Box 2 (5′-ATTGACTTGACCGTCATCGG-3′) was introduced into the upstream region of the HIS3 mini-promoter to generate the reporter construct. Additionally, yeast one-hybrid (Y1H) screening was performed to search for the putative TFs that were genetically epistatic to *VvWRKY37* (Ouwerkerk and Meijer, 2001). The promoter fragment (1,165 bp) of *VvWRKY37* was sub-cloned and introduced into the upstream region of the minimal HIS3 promoter to generate the bait vector. Total RNA was extracted from the grape vine root, shoot, leaf, and winter bud and used to construct a cDNA expression library (OE Biotech, Shanghai, China). The cDNA library screening was done by transferring the bait vector into the Y187 yeast strain. The putative transformants were identified by their growth on a medium lacking histone, leucine, and tryptophan, and containing 3-aminotriazole (3-at), (Solarbio, Beijing, China). PCR with general primer pairs was run to amplify the individual transformants

(**Supplementary Table 1**). Each PCR product was subjected to DNA sequencing, and a blasting tool from NCBI was used to verify its putative identity.

Transient Expression Assay

Partially expanded grape vine leaves were used for a transient expression assay *via A. tumefaciens* inoculation as described earlier (Wang F. et al., 2019). Histochemical staining was conducted for a GUS reporter assay after inoculation and *in vitro* co-cultivation for 3–5 days. The dual luciferase was conducted in the living tobacco leaf. The mixed suspension of *A. tumefaciens* harboring effector and reporter constructs was injected into tobacco leaves by using a single-use syringe. The reporter construct was actuated by the promoter of *PtCPY707A2*(2, 000 bp) and *VvWRKY37* (1,165 bp), respectively. The transformed tobacco plants were placed in a dark room for 1 day, followed by 3–5 days under normal growth conditions. Before determining the enzyme activity of luciferase, a tobacco leaf was cut down and sprayed with substrate solution (0.15 mg/ml D-Luciferin Sodium Salt, 0.02% Triton-X 100). After that, the leaf was kept in dark for 5 min at room temperature. Finally, a CCD device was used to visualize fluorescence images of injected leaves.

Bud Break Assay

Two-year-old poplar plants were prepared for bud break investigation. Before autumn, the transgenic and control (84 K) poplar plants were placed outside under natural environmental conditions until a bud set (at the end of October). The plants were then transferred to an incubator (4°C, dark conditions for 24 h) for 2 months to meet their chilling requirement, and then moved into a greenhouse under a long-day photoperiod (8-h dark/16-h light) at 21°C temperature. The date of bud break (emergence of green tissues at the tip of the bud) was monitored and recorded after every 2 days. The bud, together with surrounding bark, was peeled at two time points: 6 and 12 days. More than 40 buds were collected from six plants for each genotype of poplar plants. The same biological bud pools were used for measuring endogenous ABA and PA content and RT-qPCR analysis of candidate genes. The contents of ABA and PA were determined by liquid chromatography and mass spectrometry at the Institute of Chemistry, Chinese Academy of Sciences (ICCAS, Beijing).

GUS Staining

GUS activity was carried out by following a previously described protocol (Weigel and Glazebrook, 2002). Plant materials were fully immersed in a 50-ml centrifuge tube containing a staining buffer. The tubes were then kept in a vacuum pump for 10 min to remove air in the plant tissues as much as possible before incubating under dark conditions at 37°C temperature for 12 h. The staining buffer contained 0.5 mg/ml X-gluc (Yeasen, Shanghai, China), a 100-mM phosphate buffer (pH = 7.), 0.01% Triton X-100, 10-mM EDTA-Na_2, 0.5-mM $K_4Fe(CN)_6$, and 0.5-mM $K_3Fe(CN)_6$. After incubation, the reaction buffer was replaced with 75% (v/v) ethyl alcohol for 2-3 times to remove chlorophyll. Finally, the fully bleached samples were visualized using a stereoscope (Leica, Wetzlar, Germany).

Prokaryotic Expression in Recombinant Proteins

The coding sequence of VvWRKY37 without stop codon was cloned in-frame to the pET32a prokaryotic expression vectors (6HIS tag). The homologous recombination method was used to construct VvABF2-HIS and VvABI5-HIS. All gene-specific primer pairs were designed using online software CE design V1.04 (Vazyme, Nanjing, China) (**Supplementary Table 1**). The recombinant vectors were introduced into E. coli strain TransB/DE3 (TransGen, Beijing, China). Recombinant protein induction was carried out by adding 2-mM isopropyl β-D-1-thiogalactopyranoside (Solarbio, Beijing, China) in the bacterial culture at 37°C for 6 h or at 16°C overnight with constant shaking at 150 rpm. The recombinant protein was purified by ProteinIso® Ni-NTA Resinkit (TransGen, Beijing, China) according to the manufacturer's protocol. The empty prokaryotic proteins (6HIS) were also induced, purified, and then used as a control in EMSA assays.

Electrophoretic Mobility Shift Assay

The cDNA fragments, including putative W-boxes and their complementary chains, were synthesized by GENEWIZ (GENEWIZ, Suzhou, China). Each cDNA fragment was labeled with biotin by using the EMSA Probe Biotin Labeling Kit (Beyotime, Shanghai, China). The reverse complements of biotin-labeled oligonucleotides (equally mixed by volume) were annealed by a thermal cycler (Thermo Fisher, Waltham, United States) to form double-stranded DNA using the temperature gradient descent strategy. The non-biotin-labeled wildtype and mutated DNA probes acted as competitors. All primers used in this assay are shown in **Supplementary Table 1**. A 10 μl of binding reaction mixtures contained: 1-mg purified recombinant protein, a 2-μl 5- × gel shift binding buffer (50% glycerol; 50-mM $MgCl_2$; 5-mM EDTA; 10-mM DTT; 500-mM KCl; and 250-mM HEPES; pH, 7.4. A 1-μl (0.05 μM) aliquot of the labeled probe was incubated with the above-mentioned solution at 24°C for 20 min. The binding mixtures were pre-incubated with the unlabeled or mutated probe at 24°C for 20 min and then supplemented with labeled probes for another 20 min for the competition assays. The HIS protein alone was used as the negative control. After incubation, each sample was resolved using 6% non-denaturing acrylamide gel electrophoresis in a 0.5 × TBE buffer (1-mM EDTA-Na_2, 44.5-mM Tris-base, and 44.5-mM boric acid) at 100 V for 1-2 h. The gel was then transferred to a nylon membrane (GE Healthcare, London, United Kingdom) for chemiluminescent detection (Solarbio, Beijing, China).

Protein Extraction and Immunoblot Analysis

The crude protein from poplar leaf was extracted using a Plant Protein Extraction Kit (Solarbio, Beijing, China). Total protein concentration was quantified with micro-spectrophotometer

ND2000C (Thermo Fisher, Waltham, United States). The tag antibodies, anti-actin, and anti-GFP were purchased from Abmart (Abmart, Shanghai, China) and used according to the manufacturer's protocol. Immunoblot analysis was carried out by following a previously described protocol (Wang X. et al., 2019).

RESULTS

VvWRKY37 Is Specifically Expressed in Dormant Bud of Grape Vine

The expression levels of *VvWRKY* TF genes were investigated in different tissues of grapes to characterize their potential involvement in bud dormancy. Interestingly, a *VvWRKY37* exhibited higher expression levels in dormant winter buds of grape vine. We investigated the expression levels of *VvWRKY37* in winter buds at different dormancy phases: paradormancy, endodormancy, ecodormancy, bleeding period, and bud break. The *VvWRKY37* was highly expressed during dormancy establishment in the fall, and then it gradually decreased until bud break in the next spring (**Figure 1A**). Upon bud break in spring, the expression level of *VvWRKY37* was almost six times lower than that during dormancy initiation in the fall (**Figure 1A**). Tissue-specific expression analysis showed that *VvWRKY37* was expressed in roots, stems, and leaves (**Figure 1B**). The RT-qPCR result confirmed that, initially, *VvWRKY37* was highly expressed during the fruit set (7 days post-anthesis), and then it was sharply decreased from 7 dpa to 35 dpa in seeds; finally, it was barely detectable at 35 dpa (**Figure 1C**). The relative expression levels were markedly raised to 70 times and 50 times in seeds from 35 to 70 dpa (véraison berry) and from 70 to 105 dpa (mature berry), respectively (**Figure 1C**).

VvWRKY37 Responds to Low Temperature, ABA, and GA_3

Plant dormancy is highly associated with environmental and physiological factors, i.e., temperature, photoperiod, and endohormones (Liu and Sherif, 2019). To determine the regulatory role of *VvWRKY37*, the detached canes were treated with exogenous ABA, GA_3, and low temperature, respectively. Then RT-qPCR was carried out to check the change in expression levels of the *VvWRKY37* in grapevine buds. As expected, the expression level of *VvWRKY37* was significantly higher in ABA-treated cane samples taken at 24 h (h). Additionally, the expression level was 3.5- and 4.8-folds higher than that of the mock treatment at 24 h and 48 h (**Figure 2A**), respectively. Inversely, GA_3 application and low-temperature treatment attenuated the mRNA level of *VvWRKY37* in grape vine buds (**Figures 2B,C**). Under low-temperature treatment, the expression level of *VvWRKY37* decreased from 0 to 12 h and maintained a low level until 48 h (**Figure 2C**).

To investigate the promoter activity of *VvWRKY37*, the expression of the β-glucuronidase (*GUS*) reporter gene under the control of the *VvWRKY37* promoter was induced in *Arabidopsis via* Agrobacterium-mediated transformation. The positive transformants were detected by selective medium and histochemical GUS staining. The 14-day-old homozygous seedlings were used to evaluate the effects of exogenous ABA and GA_3, and low temperature on the promoter activity of *VvWRKY37*. The results of *GUS* gene expression (**Figure 2D**) and histochemical staining (**Supplementary Figure 2**) in *pVvWRKY37*:*GUS* in *Arabidopsis* showed that exogenous application of ABA induced higher *GUS* expression, while GA_3 and low temperature repressed *GUS* expression compared with the control plants.

VvWRKY37 Binds the Canonical W-Box Motif in Yeast

WRKY TFs, one of the largest gene families in plants, play dual transcriptional roles by perceiving and binding the highly conserved *cis*-element W-box TTGACC/T (Rushton et al., 2010). A short DNA sequence, defined as a BOX2 element, was firstly proved to be particularly bound by WRKY TFs (Rushton et al., 1995). To test the transcriptional activity of VvWRKY37, the bait vector was constructed by introducing a tandem-repeated BOX2 element upstream of the reporter gene. Our results showed that yeast transformants harboring the prey vector pGADT7-VvWRKY37 and the bait vector Wbox-pHIS2 grew well on a selective medium (**Figure 2E**). In contrast, no yeast cell grew on the selection medium when the core motif 5′-TGAC-3′ of the W-box element was mutated (mWbox) into 5′-caAC-3′ (**Figure 2E**). Thus, VvWRKY37 showed transcriptional activity by binding to the canonical W-box motif. However, it is still unclear whether VvWRKY37 acts as an activator or a repressor in the VvWRKY37-mediated regulatory networks.

Ectopic Overexpression of *VvWRKY37* Results in Delayed Bud Break in Poplar

We did heterologous overexpression of *VvWRKY37* in perennial deciduous poplar to explore its biological function in bud dormancy. The leaves of poplar "84K" (*Populusalba* × *Populusgllandulosa*) plants were used to generate stable transgenic plants *via* Agrobacterium-mediated transformation. More than ten independent transgenic poplars were obtained according to the result of PCR amplification (**Supplementary Figure 3**). Three lines (Line 6, Line 32, and Line 33) were chosen for further investigation, while the wildtype (WT) poplar was used as control. The immunoblot assay verified that the fusion protein VvWRKY37-GFP was substantially overexpressed in three transgenic poplar plants (**Figure 3A**). The bud break phenotypes were investigated among WT and the transgenic poplar plants after dormancy release under natural conditions. Three transgenic lines exhibited delayed bud break compared with the WT (**Figure 3B**). The initiation of the endodormancy phase was determined by the bud break dynamics assay. The results showed that heterologous overexpression of *VvWRKY37* impeded the initiation time of bud break, which led to delayed bud break for transgenic poplar plants (**Figure 3C**). The bud break initiation in transgenic poplar plants was delayed for 4 days compared to WT poplar plants (**Figure 3C**). Additionally, the delayed efficiency on winter bud break positively related

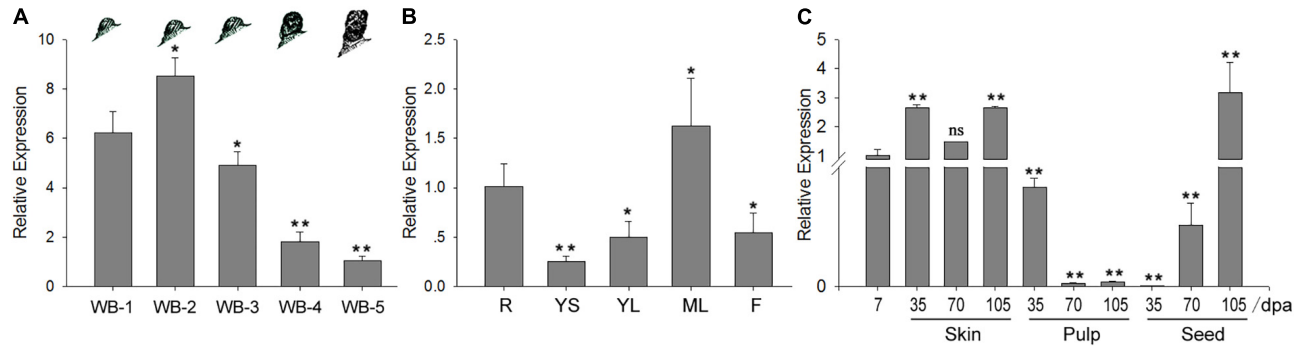


FIGURE 1 | The *VvWRKY37* expression pattern in grape vine. **(A–C)** The relative expression values of VvWRKY37 gene in different grape tissues using qRT-PCR., WB-1: winter buds at the beginning of leaf fall in autumn, WB-2: winter buds in the endodormancy phase, WB-3: winter buds during transition from endodormancy to ecodormancy, WB-4: winter buds at ecodormancy, WB-5: the green tip stage of winter buds, which was defined as bud break. R: root, YS: young stem, YL: young leaf, ML: mature leaf, I: inflorescence (full bloom). *VvActin* is used as the reference gene, and gene expression is normalized to WB-5 **(A)**, root **(B)** or 7 dpa **(C)** expression levels, which are assigned a value of 1. Each datum represents the average of triple biological replicates + SD. Asterisks in panels **(A–C)** indicate statistically significant differences compared with the level of WB-1 **(A)**, root **(B)** or 7 dpa **(C)**: $*p < 0.05$, $**p < 0.01$. ns, no significance; dpa, days post-anthesis.

to the gene expression level of *VvWRKY37* in transgenic poplar. The transgenic poplar line 32 with the strongest over-expression of *VvWRKY37* represented the latest bud break during the spring season.

Abscisic Acid Catabolism Related Gene *PtCYP707A2* Is Repressed in Transgenic Poplar Plants

Earlier, we have proved that *VvWRKY37* delayed bud break in transgenic poplar plants. However, there is no reported work available on the signaling pathway modulated by VvWRKY37 in plant dormancy-related regulatory networks. We firstly measured ABA and the catabolite of winter buds in WT and the transgenic poplar plants to better understand their correlation with VvWRKY37. Regardless of bud break status, the ABA levels in buds of transgenic poplar plants were higher than that of WT poplar plants (**Figure 4A**). In contrast, the level of the ABA catabolite phaseic acid (PA) was significantly decreased in transgenic plants compared with WT poplar plants (**Figure 4B**). We analyzed transcript levels for central components involved in ABA biosynthesis and catabolism to detect the molecular mechanism behind increased levels of ABA in the transgenic poplar plants. Our results showed that the expression level of *PtCYP707A2* was significantly downregulated in the transgenic poplar plants (**Figure 4C**). The expression level of *PtCYP707A2* intransgenic line 32 was 5-fold lower than that of the WT poplar plants. However, no change in the expression level was observed for two other paralog genes *PtCYP707A1* and *PtCYP707A4* (**Figure 4C**). Additionally, no noticeable difference was observed in the expression level of ABA biosynthesis genes, *PtNCED1*, *PtNCED3*, *PtNCED5*, and *PtNCED6* between the WT and transgenic poplars plants (**Supplementary Figure 4**).

The above-mentioned findings suggested that *VvWRKY37* possibly enhanced ABA accumulation *via* transcriptional repression of ABA catabolic genes. Firstly, we explored available evidence to verify whether VvWRKY37 directly regulated *PtCYP707A2*. The results of the transcriptional activity assay showed that VvWRKY37 could bind to the W-box *cis*-element (**Figure 2E**). The *cis*-regulatory element analysis in the upstream promoter region of the *PtCYP707A2* gene resulted in identification of three candidate W-box elements. Each of the W-box elements, together with the flanking sequence, was designed as a DNA probe and named as PtA2-P1, PtA2-P2, and PtA2-P3, respectively (**Figure 4D** and **Supplementary Figure 5A**). The recombinant protein VvWRKY37-HIS was prepared by using a prokaryotic expression method. The electrophoretic mobility shift assay (EMSA) results indicated that VvWRKY37 specifically interacted with the DNA probe PtA2-P2, and the competitive assay further confirmed the protein-DNA interaction *in vitro* (**Figure 4E** and **Supplementary Figure 5B**). The core motif of the W-box was irreplaceable for the proper interplay between VvWRKY37 and DNA probe PtA2-P2 (**Figure 4E**). The EMSA results were cross-checked by using the Y1H assay (**Figure 4F**). The yeast cells harboring VvWRKY37 and wildtype W-box motifs grew well on the selective medium, while the negative control could not grow (**Figure 4G**). The VvWRKY37 failed to activate the reporter gene in yeast cells when the W-box motif 5′-CATTGACG-3′ was mutated into 5′-aATcaACG-3′ (**Figure 4G**). The transient expression assay showed that heterologous overexpression of VvWRKY37 in tobacco leaves decreased luciferase gene expression initiated by the promoter of *PtCYP707A2* (**Figures 4H,I**). Co-transforming *35S:VvWRKY37* with *pPtCYP707A2:LUC* significantly attenuated luciferase activity compared with the empty vector (**Figures 4H,I**). The VvWRKY37 did not depress the expression of the luciferase gene initiated by the mutant promoter of *PtCYP707A2* on the deletion of W-box from it, indicating that the W-box motif was indispensable for VvWRKY37 modulating the promoter activity of *PtCYP707A2* (**Supplementary Figure 6**). Therefore, from the above analysis, we inferred that VvWRKY37 can directly bind to the promoter

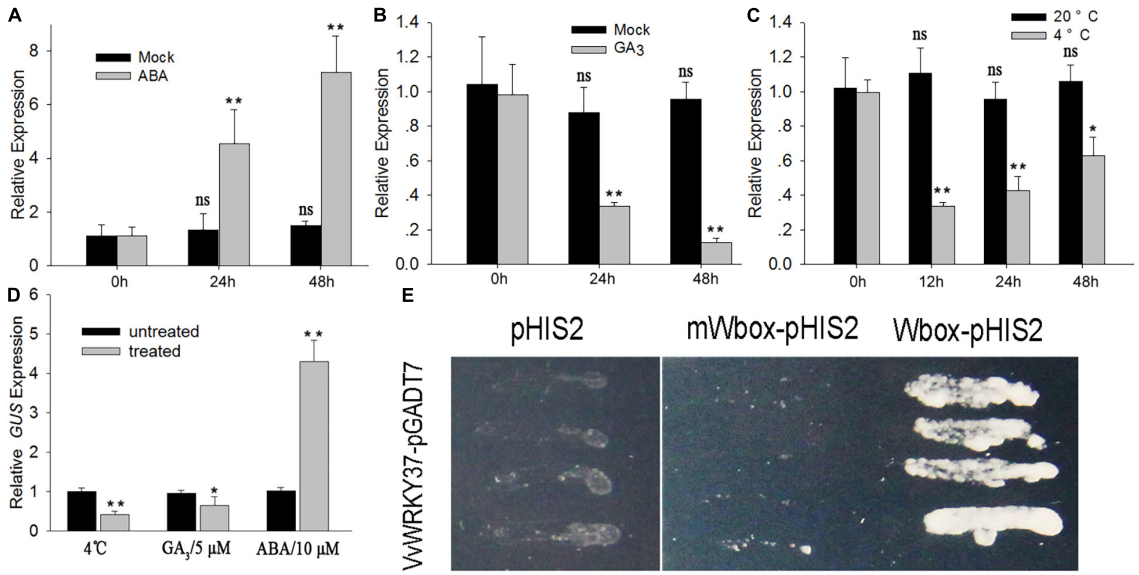


FIGURE 2 | *VvWRKY37* expression responds to exogenous phytohormones and low-temperature treatments. **(A–C)** The relative expression value of *VvWRKY37* in grape buds responding to ABA (10 μM), GA_3 (5 μM), and low temperature (4°C) treatments. Detached grape vine canes were sampled for different treatments at the beginning of leaf fall. Mock treatment is the control. *VvActin* is used as an internal control, and the default expression value was 1 at 0 h after treatment. **(D)** The transcript abundance of *GUS* gene in the leaf of *pVvWRKY37*:*GUS*. The 14-day-old seedlings were treated with ABA (10 μM), GA_3 (5 μM), and low temperature (4°C) for 24 h. The 5^{th} and 6^{th} leaves were harvested for RNA extraction. *AtGAPDH* is used as an internal control, and the default expression value was 1 for control seedling. Each datum represents the average of triple biological replicates + SD. **(E)** Transcriptional activity of VvWRKY37 on the tandem-repeated BOX2 element, which contains four W-box *cis*-elements. The core motif sequence 5′-TGAC-3′ of W-box was mutated into caAC in m Wbox. The selective medium is nutrition deficiency in histone, leucine, and tryptophan. Asterisks in panels **(A–D)** indicate statistically significant differences compared with 0 h: $^*p < 0.05$, $^{**}p < 0.01$. ns means no significance.

region of *PtCYP707A2* and transcriptionally repressed the promoter activity.

VvWRKY37 Binds to the Promoter of *VvCYP707A4.2* and Suppresses Its Expression in Grape Vine

The expression of *VvWRKY37* was induced in response to the exogenous application of ABA (**Figure 2A**), which implied the potential relationship between *VvWRKY37* and ABA levels in the grape vine. The ABA contents of grape vine buds were measured from paradormancy to bud break. The data presented in **Figure 5A** suggested that the highest ABA level was found at the time of bud dormancy establishment in the winter. The ABA content was reduced in grape vine winter buds from the endodormancy phase to bud break. It was likely that the expression level of *VvWRKY37* was directly proportional to the ABA level during bud dormancy periods (**Figures 1A**, **5A**). The RT–qPCR results illustrated that relative expressions of *VvCYP707A2*, *4.2*, and *4.3* exhibited dramatic changes in grape vine winter buds from dormancy establishment to bud break (**Figures 5B–D**). The winter bud presented a relatively low relative expression level of *VvCYP707A2*, *4.2*, and *4.3* from dormancy establishment (WB-1) to the endodormancy phase (WB-2), while it sharply increased after the endodormancy phase (WB-3) (**Figures 5B–D**). From the end of the endodormancy phase to bud break, the relative expression of *VvCYP707A2*, *4.2*, and *4.3* was downregulated to an extremely low level (**Figures 5B–D**). The relative expression levels of *VvCYP707A2*, *4.2*, and *4.3* reached to a peak at the end of the endodormancy phase.

We carried out some additional assays to verify the molecular relationship between VvWRKY37 and *VvCYP707A4.2*. The binding affinity between VvWRKY37 and W-box *cis*-element adjacent to the transcription start site of *VvCYP707A4.2* was confirmed by both EMSA and Y1H assays (**Figures 5E–H**). The EMSA assay results showed that VvWRKY37 could not bind to the other two W-box motifs (**Supplementary Figures 7A,B**). Interestingly, we found that the core motif of W-box *cis*-element (5′-CATTGACG/T-3′) bound by VvWRKY37 was similar to the core motif (5′-GATTGACT-3′) in tandem-repeated Box 2, which was tested earlier in the transcription activity assay (**Figure 2E**). We failed to find a similar motif in the promoter regions of *PtCYP707A1*, *PtCYP707A4*, *VvCYP707A2*, and *VvCYP707A4.3* (**Supplementary Table 2**). Preliminary evidence supported our argument that 5′-G/CATTGACG/T-3′ might be the conserved binding preference for VvWRKY37. A transient expression assay was conducted to verify the inhibitory effect of VvWRKY37 on the W-box motifs in the promoter region of *VvCYP707A4.2* (**Figure 5I**). The reporter construct contained a tandemly repeated W-box motif of *VvCYP707A4.2*, which was inserted in the front of the mini CaMV35S promoter. The recombinant plasmid *35S:VvWRKY37-GFP* and the reporter construct Vv4.2-dP1-*minip35S:GUS* significantly attenuated the accumulation of

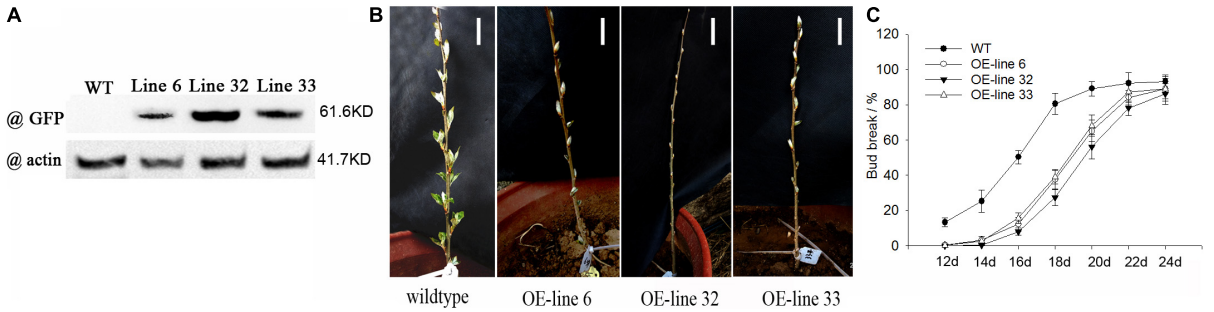


FIGURE 3 | The bud break phenotypes of transgenic plants of VvWRKY37 over-expressing in poplar. **(A)** The level of the VvWRKY37-GFP fusion protein in 35S:VvWRKY37-GFP transgenic poplar lines; an anti-GFP antibody was used in the immunoblot analysis. The anti-actin antibody was used as loading control. The numbers present the protein molecular size. **(B)** Bud-break phenotypes of WT and transgenic poplar plants in spring under environmental conditions. The scale bar is 5 cm. **(C)** The dynamic of bud break in VvWRKY37-overexpressing and control poplars over days after dormancy release. Abscissa is the time (day) after moving the poplars into the greenhouse. The values are the average of means of triple biological replicates of three transgenic lines and for five WT plants. Bars represent SD.

GUS protein in grape vine leaves compared with the empty vector *35S*:*GFP* (**Figure 5J**). The RT-qPCR results also confirmed that VvWRKY37 noticeably depressed the expression of reporter gene *GUS* (**Figure 5K**). The expression analysis results for transformed grape vine leaves showed that overexpression of *VvWRKY37* significantly downregulated the expression levels of *VvCYP707A4.2*, while the expression level of *VvCYP707A4.2* was barely altered in grape vine leaves transformed by the empty vector *35S*:*GFP* (**Supplementary Figures 8A,B**).

VvWRKY37 Is Transcriptionally Modulated by the Upstream Regulators of the ABA Signaling Pathway

ABA plays a dominant role in plant dormancy regulatory network signaling cascades through ABA-related regulators. Our results provided substantial evidence to prove that the higher levels of ABA evoked the expression of *VvWRKY37*, while VvWRKY37 inversely affected ABA homeostasis through a feedback loop. However, the mechanism by which the ABA alters the expression level of *VvWRKY37* is still unknown. There could be special factors that regulated ABA homeostasis and transduced downstream signals to the TF VvWRKY37. Firstly, we investigated the *cis*-elements in the promoter region of *VvWRKY37* using an online database (Lescot et al., 2002). Several ABRE (ABA-responsive elements) motifs were identified, implying that *VvWRKY37* was potentially regulated by ABFs (ABRE-binding factors). The ABFs, A subfamily of bZIP TFs induced by exogenous application of ABA, were positively involved in the ABA signal transduction cascade (Kim et al., 2004). We performed the Y1H library screening assay to determine whether VvWRKY37 was modulated by bZIP protein or other TFs. A DNA fragment of the *VvWRKY37* promoter was cloned in-frame to the bait vector, which was then co-transformed with grape vine mRNA library vectors. Finally, we identified a group of ripening-related proteins and an arm repeat protein (**Supplementary Table 3**). The arm repeat protein has been previously reported to interact with ABF2 and modulate the transcriptional activity of ABF2 (Kim et al., 2004). Thus, VvABF2 and its homeotic gene VvABI5 were considered intermediates that linked the ABA signal and *VvWRKY37* mRNA abundance in the dormancy regulatory network. A series of analyses were carried out to find enough evidence in support of the upstream regulator of *VvWRKY37*.

The expression level of *VvABF2* and *VvABI5* reached to peak at the bud dormancy establishment stage and then continuously declined to a low level until the bleeding period (**Figures 6A,B**), which was almost in parallel to the change of *VvWRKY37*expression levels (**Figure 1C**). To conduct the EMSA assay, ABRE elements in the promoter region of *VvWRKY37* was designed as DNA probe W37-P1 and W37-P2 (**Figure 6C**). Both VvABF2 and VvABI5 genes were capable of binding with biotin-labeled DNA fragment W37-P2 but failed to bind with W37-P1 (**Figures 6D,E** and **Supplementary Figure 9**). The cold mutant DNA probe Cold-mP2 could not competitively subdue the shifted protein-DNA complex, implying a distinctive binding activity of VvABF2 and VvABI5. Generally, VvABF2 and VvABI5 were considered inducible by the ABA signal and transcriptionally provoked the expression of their downstream target *VvWRKY37*. The transient expression assay confirmed that heterologous overexpression of VvABF2 or VvABI5 in tobacco leaves elicited an enhanced expression of the luciferase reporter gene initiated by the promoter of *VvWRKY37* (**Figures 6F,H**). Co-transformation of *35S*:*VvABF2* or *35S*:*VvABI5*, together with *pVvWRKY37*:*LUC*, markedly increased luciferase activity than that of the empty effector vector in tobacco leaves (**Figures 6G,I**). Both VvABF2 and VvABI5 did not show an inhibitory effect on the luciferase activity when replacing the ABRE motif of the *VvWRKY37* promoter (**Supplementary Figures 10A,B**). The result indicated that the ABRE motif was indispensable for proper interplay between VvABI5 or VvABF2 and *VvWRKY37*.

We also tested the transcriptional activity of VvABF2 and VvABI5 in grape vine leaves *via* the *Agrobacterium*-mediated transformation method. The schematic diagrams of the reporter and effector constructs are shown in **Figure 6J**. The overexpression of VvABF2 or VvABI5 stimulated a strong expression of the *GUS* reporter gene initiated by the *VvWRKY37* promoter (**Figures 6K,L**). It seemed that the ABA-inducible TFs,

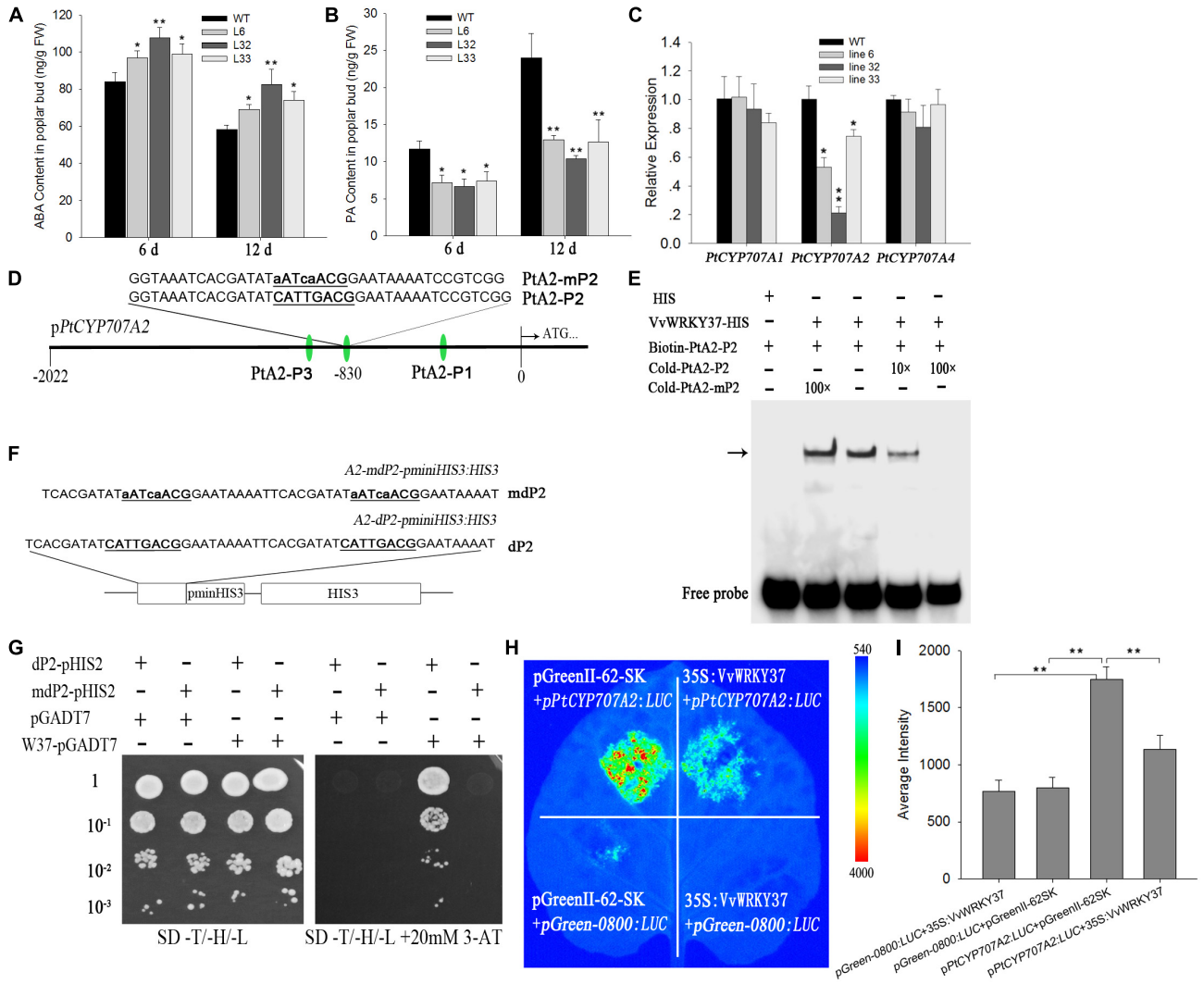


FIGURE 4 | VvWRKY37 affects abscisic acid catabolism. **(A,B)** ABA and PA content in transgenic and WT poplars. Fresh buds were accurately weighed for metabolite quantification. Each datum represents the average of triple biological replicates + SD. **(C)** Expression patterns of *PtCYP707A*s in *35S:VvWRKY37* and WT poplars. *PtUbi* is used as an internal control, and gene expression is normalized to the WT expression level, which is assigned a value of 1. Each datum represents the average of triple biological replicates, and error bars indicate SD. **(D)** Nucleotide sequences of the probes derived from the *PtCYP707A2* promoter are used for VvWRKY37-His recombinant protein-binding assays. The core sequence 5′-CATTGACG-3′ of PtA2-P2 is mutated into 5′-aATcaACG-3′ in PtA2-mP2. **(E)** DNA-binding affinity of VvWRKY37 to W-box fragment PtA2-P2. Competition assays were done with an increasing amount of an unlabeled probe (10-fold and 100-fold). The unlabeled probe mP1 is 100-fold excess of labeled probe P1 in Lane 2 (from left to right). The position of protein-DNA complexes is indicated by an arrow. **(F,G)** VvWRKY37 can bind to the W-box *cis*-element of the *PtCYP707A2* promoter *in vitro* by a yeast one-hybrid assay. The nucleotide fragment A2-dP2 contains two tandem repeats of 5′-CATTGACG-3′, which are both mutated into 5′-aATcaACG-3′ in A2-mdP2. 10^{-1}, 10^{-2}, and 10^{-3} indicate the yeast concentration diluted with a 0.9% NaCl buffer by 0.1, 0.01, and 0.001. 3-AT stands for 3-Amino-1, 2, 4-triazole. **(H)** Luciferase assays were conducted *via* co-injecting *A. tumifaciens* into *Nicotianabenthamiana* leaves. The activities of luciferase were imaged using a CCD device. **(I)** The average intensity of fluorescence exhibited on the tobacco leaves. The fluorescence values were calculated using the software Image J (1.46 R). Asterisks in panels **(A–C,I)** indicate statistically significant differences as indicated: $^{*}p < 0.05$, $^{**}p < 0.01$.

VvABF2, and VvABI5 transcriptionally activated the promoter of *VvWRKY37* through direct protein-DNA interactions. The grape vine leaves inoculated with *pVvWRKY37*:*GUS* accumulated much more GUS protein than under ABA application (**Figures 6K,L**). The RT-qPCR assay was used to measure the changes in transformed grape vine leaves. Overexpression of VvABF2 or VvABI5 increased the expression level of *VvWRKY37* in grape vine leaves (**Supplementary Figures 11A–C**). Likewise, the expression level of *VvWRKY37* in transformed leaves was significantly induced by ABA application.

DISCUSSION

Plant WRKY TFs, one of the largest TF families, are increasingly attracting researchers to explore their novel functions, including

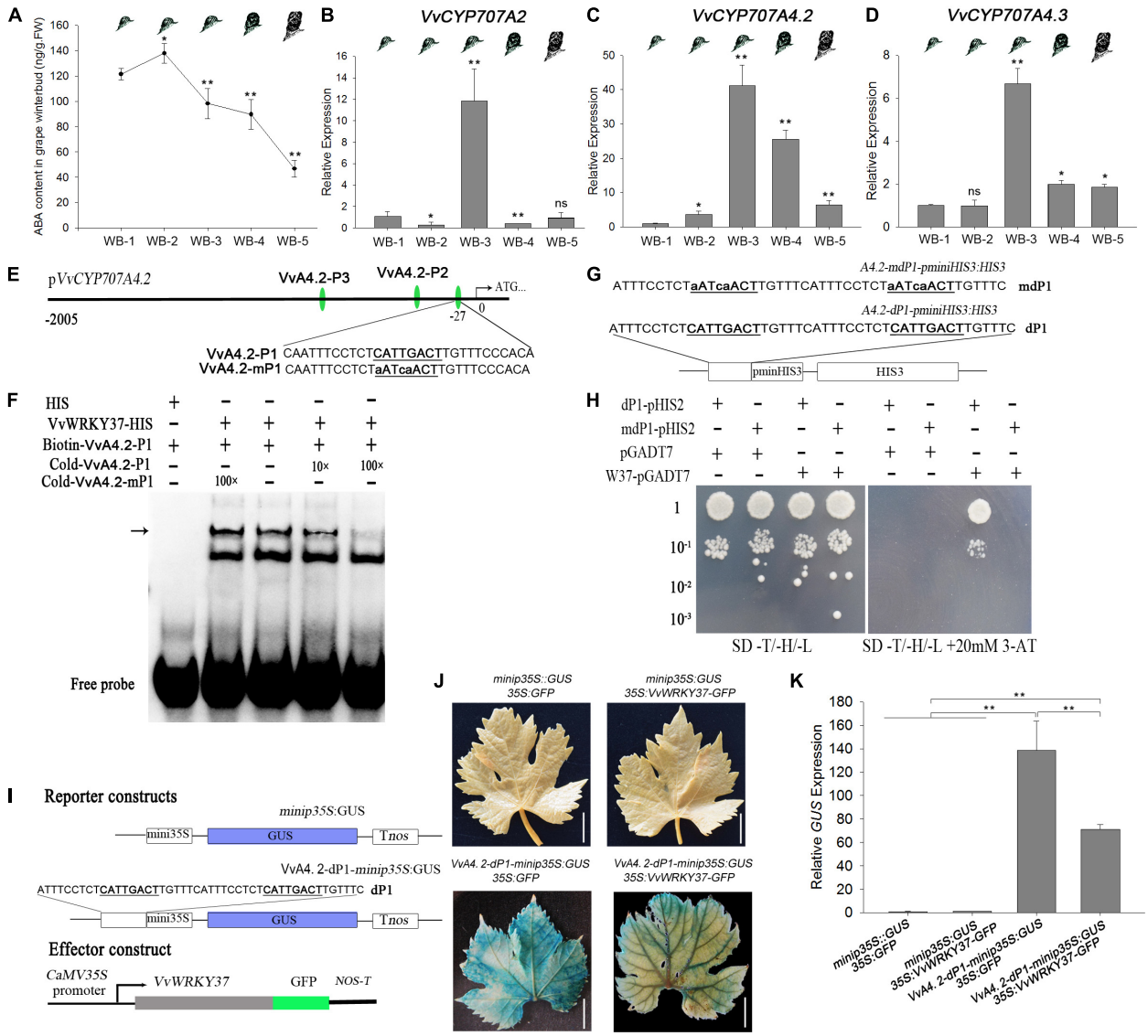


FIGURE 5 | VvWRKY37 modulates the expression of *VvCYP7074.2* in grape vine. **(A)** ABA content in grape vine winter buds at different dormancy phases. Data are means of five biological replicates ± SD. **(B–D)** *VvCYP707A2*, *4.2*, and *4.3* expressions in grape winter buds at different dormancy phases. *VvActin* serves as the reference gene, and the mRNA abundance is normalized to the WB-1-stage expression level assigned a value of 1. Each datum represents the average of triple biological replicates + SD. **(E)** Nucleotide sequences of the DNA probes derived from *VvCYP707A4.2* promoter. The W-box *cis*-element 5′-CATTGACT-3′ of VvA4.2-P1 is mutated into 5′-aATcaACT-3′ in VvA4.2-mP1. **(F)** The EMSA assay was employed to test DNA-binding affinity of VvWRKY37 to the W-box fragments VvA4.2-P1. Competition assays were performed just as described in **Figure 4E**. The position of protein-DNA complexes is indicated by an arrow. **(G,H)** VvWRKY37 can bind to the W-box *cis*-element of the *VvCYP707A4.2* promoter *in vitro* by a yeast one-hybrid assay. The nucleotide fragment A4.2-dP1 contains two tandem repeats of 5′-CATTGACT-3′, which are both mutated into 5′-aATcaACT-3′ in A4.2-mdP1. 10^{-1}, 10^{-2}, and 10^{-3} indicate the yeast concentration diluted with a 0.9% NaCl buffer by 0.1, 0.01, and 0.001. 3-AT stands for 3-Amino-1, 2, 4-triazole. **(I)** Schematic diagrams of the reporter and effector constructs used in the transient co-transformation of grape vine leaves *via A. tumifaciens*. The nucleotide fragment A4.2–dP1 contains two tandem repeats of W-box *cis*-element 5′-CATTGACT-3′, which are both mutated into 5′-aATcaACT-3′ in A4.2–mdP1. **(J)** The images of GUS-stained grape vine leaves after injecting *A. tumifaciens*. All white scale bars indicate 1 cm. **(K)** The transcript level of *GUS* gene in transformed grape vine leaves. Asterisks in panels **(A–D,K)** indicate statistically significant differences as indicated: $^{*}p < 0.05$; $^{**}p < 0.01$. ns means no significance compared with WB-1.

phytohormone-related biological processes (Rushton et al., 2012; Ding et al., 2014; Liu et al., 2015; Chen et al., 2016; Huang et al., 2016; Zhou et al., 2020). Plant bud dormancy induction, duration, and release were directly associated with environmental conditions, photoperiod, temperature, and endogenous phytohormones (Liu and Sherif, 2019). Genome-wide expression analysis and comparative RNA-Seq analysis have indicated the regulatory role of several *WRKY* genes underlying bud dormancy in perennial plants (Chen et al., 2016; Khalil-Ur-Rehman et al., 2017). However, there is no reported knowledge

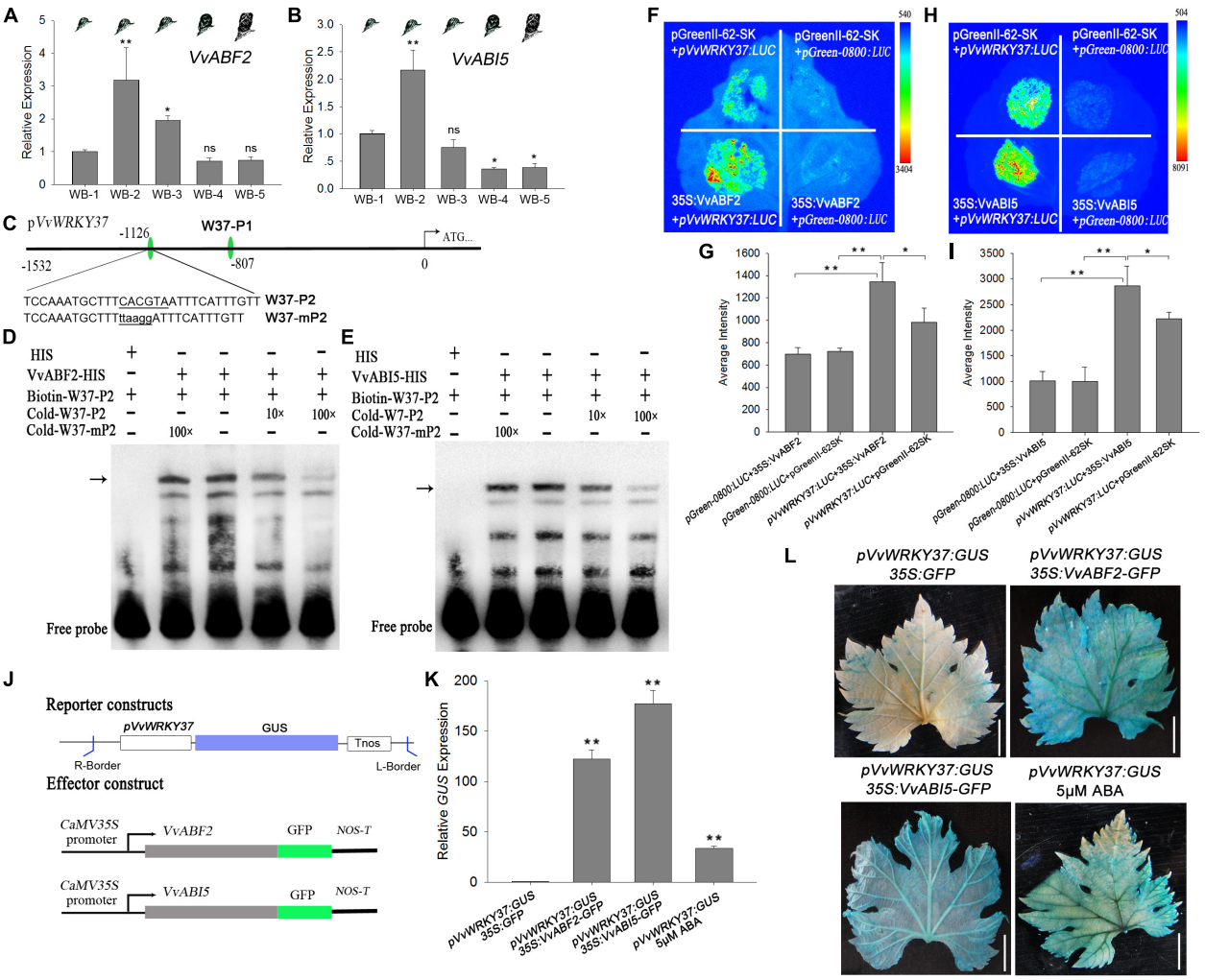


FIGURE 6 | VvWRKY37 is implicated in ABA cascade response *via* upstream regulator VvABF2 and VvABI5. **(A,B)** The mRNA levels of *VvABF2* and *VvABI5* in grape winter buds at different dormancy phases. *VvActin* is used as an internal control, and gene expression is normalized to WB-1-stage expression levels, which were both assigned a value of 1. Each datum represents the average of triple biological replicates + SD. **(C)** Nucleotide sequences of the DNA probes W37-P2 derived from *VvWRKY37* promoter were used for EMSA assays. W37-P2 contains a putative ABRE *cis*-element 5′-CACGTA-3′, which is mutated into 5′-ttaagg-3′ in W37-mP2. **(D,E)** Electrophoretic mobility shift assays for detecting the binding affinity of VvABF2-His **(D)** and VvABI5-His **(E)** fusion proteins. Competition assays were performed just as described in **Figure 4E**. The position of protein-DNA complexes is indicated by a black arrow. **(F,H)** Luciferase assays were conducted in *Nicotianabenthamiana* leaves *via A. tumifaciens* as described in **Figure 4H**. The full length promoter of *VvWRKY37* was used in the assays. **(G,I)** The average intensity of fluorescence exhibited on the tobacco leaves. All data are shown as the average of triple biological replicates + SD. **(J)** Schematic diagrams of the reporter and effector constructs for transient expression in grape vine leaves. **(K)** The relative expression of GUS gene in grape vine leaves; *VvActin* is used as an internal control. Each datum represents the average of triple biological replicates + SD. **(L)** Histochemical staining for detecting GUS protein expression in grape vine leaves. Scale bars = 1 cm. Asterisks in panels **(A,B,G,I,K)** indicate statistically significant differences as indicated: $*p < 0.05$, $**p < 0.01$. ns means no significance compared with WB-1.

available on intricate molecular regulatory mechanisms for WRKY TF in bud break. Here, we reported *VvWRKY37*, which was highly expressed in the dormant winter bud of the grape vine. Ectopic overexpression of *VvWRKY37* delayed bud break in transgenic poplar plants. A series of experiments provided a new perspective in which VvWRKY37 modulated bud break through ABA-mediated signaling cascades. Functional validation of VvWRKY37 built a bridge between phytohormone ABA and downstream gene function in plant dormancy regulatory cascade (**Figure 7**).

Previously, 20 *WRKY* genes related to bud dormancy have been identified by transcriptomic approach from different dormancy phases, summer buds, and winter buds during paradormancy and endodormancy (Khalil-Ur-Rehman et al., 2017). Among these differentially expressed *VvWRKY*s, the expression level of *VvWRKY37* (LOC100267688) decreased at the onset of the endodormancy phase in grape vine buds (Khalil-Ur-Rehman et al., 2017). Coincidently, the expression level of *VvWRKY37* declined continuously from endodormancy to bud break in grape vine winter buds (**Figure 1A**). The

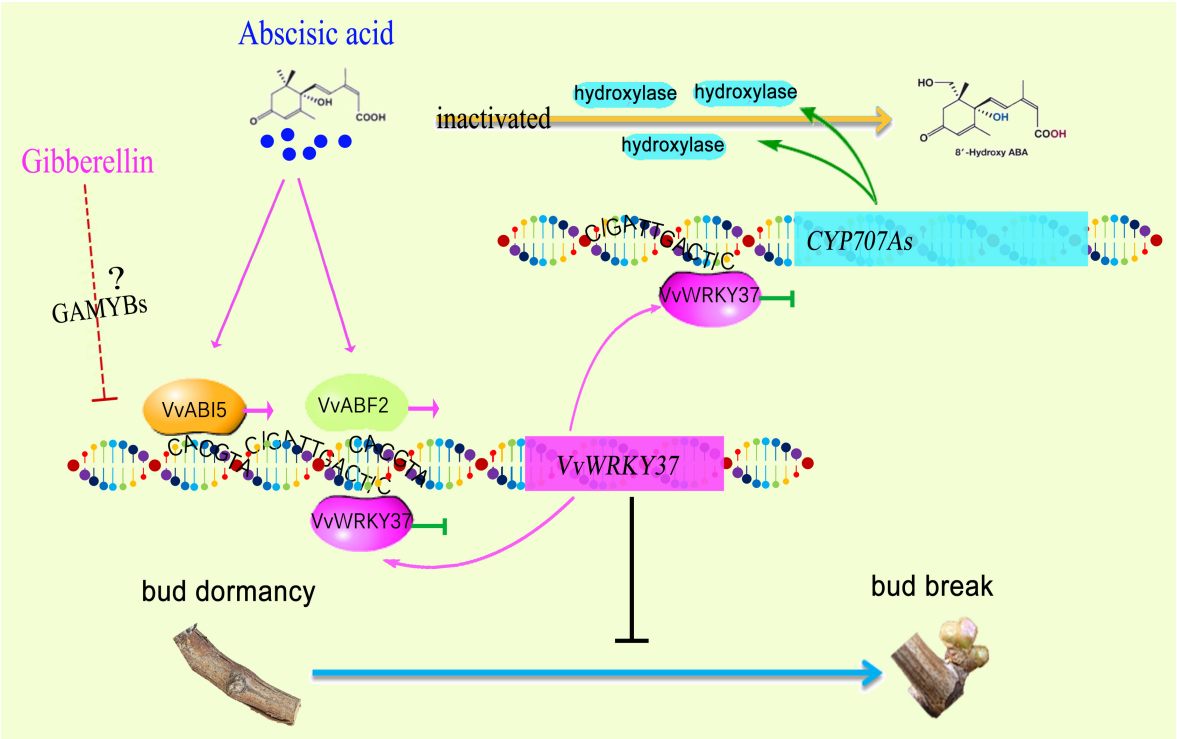


FIGURE 7 | A proposed model for the role of VvWRKY37 in ABA-mediated dormancy regulatory cascades.

dominant roles of ABA in modulating plant dormancy have been reported in numerous former pieces of literature (Tylewicz et al., 2018; Liu and Sherif, 2019). For an instance, exogenous application of ABA has inhibitory effects on the dormancy release of grape vine buds (Zheng et al., 2015. Typically, the ABA content of grape vine buds increases prior to the endodormancy phase at the onset of autumn (Rubio et al., 2019). Similarly, higher ABA levels are vitally important in avoiding viviparous seeds before release from the mother plant (Kushiro et al., 2004). Our results showed that change in the ABA content was directly proportional to expression levels of *VvWRKY37* in grape vine buds (**Figures 1A**, **5A**), implying an underlying regulatory mechanism. Additionally, we have provided evidence for induction of *VvWRKY37* gene under exogenous application of ABA, and its regulatory role in inversely modulated ABA homeostasis by repressing the catabolism of rate-limiting enzymatic genes *CYP707A*. ABA metabolism is an essential pathway in the ABA-mediated regulation of plant dormancy (Zheng et al., 2015). Overexpression of *VvA8H-CYP707A4* in grape vine reduced the ABA level and enhanced bud dormancy release (Zheng et al., 2018a). Previously, it has been reported that the *NCED* and *ABA8′OH* catalytic genes that determine ABA homeostasis are transcriptionally regulated by WRKY TFs (Liu et al., 2015; Luo et al., 2017). In the current study, the expression patterns of *WRKY37* and *CYP707As* were not highly correlated in grape vine winter buds (**Figures 1A**, **5B–D**). At least, WRKY37-mediated transcriptional repression of *CYP707As* was not found from the bleeding period to bud break. All their transcript abundances were relatively low during this stage. The actual regulatory mechanism might be more complex than our current findings. However, we provided direct evidence to prove that *VvWRKY37* enhanced ABA accumulation *via* repressing the ABA catabolic pathway, but not regulating ABA biosynthesis.

WRKY TFs were characterized for recognizing and binding the *cis*-element 5′-TTGACT/C-3′ (W-box) with an invariant core motif 5′-TGAC-3′ of downstream targets (Rushton et al., 2010). However, the adjacent DNA sequences outside the core motif are sometimes conclusive for the specificity of WRKY-binding affinity (Ciolkowski et al., 2008). The flanking nucleotide is crucial for individual WRKY TF to determine which W-box it should specifically bind in the complex context (Ciolkowski et al., 2008). The binding affinity of the fused protein VvWRKY37-HIS to various DNA probes with different flanking sequences was widely investigated *via* EMSA assays. The VvWRKY37 exhibited a binding affinity to the 5′-G/CATTGACT/C/G-3′*cis*-element (**Supplementary Figure 12**). Notably, this binding behavior occurred not only in grape vine but also in heterologous plant species, i.e., poplar. It has been speculated that the preferential binding affinity of VvWRKY37 could have originated from lower plant species. Furthermore, dissecting the binding preference of VvWRKY37 could identify the downstream targets involved in VvWRKY37-like TFs-mediated regulatory pathways.

The phytohormone GA promotes seed germination (Née et al., 2017), but its effect on bud break is not absolutely explicit (Zheng et al., 2018b). Although we did not pay

much attention to the roles of VvWRKY37 in GA-mediated or chilling-related dormancy regulation, our results imply that VvWRKY37 might be involved in the above-mentioned regulatory cascades. The exogenous application of GA and low-temperature exposure reduced the expression level of *VvWRKY37* in the grape vine (**Figures 2B,C**). The GA and low temperature acted as bud dormancy release stimulators by directly inhibiting the expression of *VvWRKY37*, or by antagonizing ABA biosynthesis or/and signaling (Seo et al., 2006). However, it was inconsistent that the GA_1 level in grape vine buds declined from dormancy induction to maintenance (Zheng et al., 2018b), while *VvWRKY37* was downregulated. Perhaps, ABA played a predominant role in modulating VvWRKY37 expression during dormancy initiation and duration, which dissembled the inhibitory effect induced by GA. During bud break, the content of ABA was stable at a low level, while increasing GA-depressed *VvWRKY37* expression. Thereby, the expression level of *VvWRKY37* was downregulated with decreasing ABA levels in grape vine buds before bud break. A previous study has reported that ABA and GA had an antagonistic effect on the regulation of several biological processes (Gazzarrini et al., 2004). Thus, the interplay between the hormones determines their effectiveness.

ABA modulated the expression of *VvWRKY37*, possibly *via* the ABA-stimulated expression of ABI5 and ABF2 (**Figure 6**). Notably, ABA exhibited a traditional signaling transduction cascade response from signaling molecules to downstream functional genes. Despite the monodirectional top-to-bottom signaling cascade, several secondary regulatory strategies, i.e., the feedback loop, synergistically made the signaling pathway work optimally, especially for hormone cascade responses (Fukazawa et al., 2017; Wang X. et al., 2019). Here, we found that high levels of ABA-induced *VvWRKY37* expression and, in turn, VvWRKY37 feedback enhanced ABA accumulation *via* transcriptional repression of the ABA catabolic gene. It was a positive feedback loop that maintained the triggered role of ABA in *VvWRKY37* expression. However, no unique W-box *cis*-element bound by VvWRKY37 was found in the promoters of ABI5 and ABF2, indicating a monodirectional relationship between VvABI5 or VvABF2 and *VvWRKY37*. VvABI5 and VvABF2 might have redundant functions in regulating *VvWRKY37* because their binding sites on the promoter of *VvWRKY37* were identical, and both of them activated the *VvWRKY37* promoter.

As an effective regulatory pathway, a homeostasis mechanism must be in place to avoid the inhibitory effect of VvWRKY37 on target genes. Previously, WRKY TF has been shown to be transcriptionally self-regulated *via* direct protein-DNA interaction in *Arabidopsis* (Mao et al., 2011). Two adjacent W-box *cis*-elements, 5′-CATTGACC-3′ and 5′-CATTGACT-3′, were identified in the promoter of *VvWRKY37*. As expected, VvWRKY37 showed high-binding affinity to both W-box motifs, which triggered a suppression effect on the reporter gene (**Supplementary Figure 13**). These results suggest that the expression level of *VvWRKY37* could be modulated *via* a negative feedback loop. These fine-tuned feedback regulations evoked by VvWRKY37 increased our understanding of the complex dormancy-related signal transduction cascades. Overall, we comprehensively dissected the transcriptional roles of VvWRKY37 in the ABA-mediated bud dormancy signal transduction cascades, which unravel novel insights into the molecular regulatory mechanisms underlying different plant dormancy phases.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

F-PW designed the study, conducted the experiments, analysis and interpretation of the data, drafted or revised the article. P-PZ conducted the experiments, analysis, and interpretation of the data. LZ revised the article. HZ conceptualized and designed the study. MA revised the article. Y-PD conceived and designed the study and drafted or revised the article. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by grants from the National Key Research and Development Program of China (2019YFD1000101) and the National Natural Science Foundation of China in 2012 (31201609).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.929892/full#supplementary-material

REFERENCES

Artlip, T., Mcdermaid, A., Ma, Q., and Wisniewski, M. (2019). Differential gene expression in non-transgenic and transgenic "M.26" apple overexpressing a peach CBF gene during the transition from eco-dormancy to bud break. *Hortic. Res.* 6:86. doi: 10.1038/s41438-019-0168-9

Busov, V., Carneros, E., and Yakovlev, I. (2016). EARLY BUD-BREAK1 (EBB1) defines a conserved mechanism for control of bud-break in woody perennials. *Plant Signal. Behav.* 11:e1073873. doi: 10.1080/15592324.2015.1073873

Chen, M., Tan, Q., Sun, M., Li, D., Fu, X., Chen, X., et al. (2016). Genome-wide identification of WRKY family genes in peach and analysis of WRKY expression during bud dormancy. *Mol. Genet. Genomics* 291, 1319–1332. doi: 10.1007/s00438-016-1171-6

Chmielewski, F. M., Gotz, K. P., Homann, T., Huschek, G., and Rawel, H. M. (2017). Identification of endodormancyrelease for cherries (*PrunusAvium L.*) by abscisicacid and sugars. *J. Hortic.* 4:3. doi: 10.4172/2376-0354.1000210

Ciolkowski, I., Wanke, D., Birkenbihl, R. P., and Somssich, I. E. (2008). Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues

into WRKY-domain function. *Plant Mol. Biol.* 68, 81–92. doi: 10.1007/s11103-008-9353-1

Ding, Z. J., Yan, J. Y., Li, G. X., Wu, Z. C., Zhang, S. Q., and Zheng, S. J. (2014). WRKY41 controls *Arabidopsis* seed dormancy *via* direct regulation of ABI3 transcript levels not downstream of ABA. *Plant J.* 79, 810–823. doi: 10.1111/tpj.12597

Fukazawa, J., Mori, M., Watanabe, S., Miyamoto, C., Ito, T., and Takahashi, Y. (2017). DELLA-GAF1 complex is a main component in gibberellin feedback regulation of GA20 Oxidase 2. *Plant Physiol.* 175:1395. doi: 10.1104/pp.17.00282

Gambino, G., Perrone, I., and Gribaudo, I. (2008). A rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochem. Anal.* 19, 520–525. doi: 10.1002/pca.1078

Gazzarrini, S., Tsuchiya, Y., Lumba, S., Okamoto, M., and Mccourt, P. (2004). The transcription factor FUSCA3 controls developmental timing in *Arabidopsis* through the hormones gibberellin and abscisicacid. *Dev. Cell* 7, 373–385. doi: 10.1016/j.devcel.2004.06.017

Huang, Y., Feng, C., Ye, Q., Wu, W., and Chen, Y. (2016). *Arabidopsis* WRKY6 transcription factor acts as a positive regulator of abscisicacid signaling during seed germination and early seedling development. *PLoS Genet.* 12:e1005833. doi: 10.1371/journal.pgen.1005833

Khalil-Ur-Rehman, M., Sun, L., Li, C. X., Faheem, M., Wang, W., and Tao, J. M. (2017). Comparative RNA-seq based transcriptomic analysis of bud dormancy in grape. *BMC Plant Biol.* 17:18. doi: 10.1186/s12870-016-0960-8

Kim, S., Choi, H. I., Ryu, H. J., Park, J. H., Kim, M. D., and Kim, S. Y. (2004). ARIA, an *Arabidopsis*arm repeat protein interacting with a transcriptional regulator of abscisicacid-responsive gene expression, is a novel abscisicacid signaling component. *Plant Physiol.* 136, 3639–3648. doi: 10.1104/pp.104.049189

Kushiro, T., Okamoto, M., Nakabayashi, K., Yamagishi, K., Kitamura, S., Asami, T., et al. (2004). The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8'-hydroxylases: key enzymes in ABA catabolism. *EMBO J.* 23, 1647–1656. doi: 10.1038/sj.emboj.7600121

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, J., Yan, X., Yang, Q., Ma, Y., Yang, B., Tian, J., et al. (2019). PpCBFs selectively regulate PpDAMs and contribute to the pear bud endodormancy process. *Plant Mol. Biol.* 99, 575–586. doi: 10.1007/s11103-019-00837-7

Liu, J., and Sherif, S. (2019). Hormonal orchestration of bud dormancy cycle in deciduous woody perennials. *Front. Plant Sci.* 10:1136. doi: 10.3389/fpls.2019.01136

Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R. P., and Somssich, I. E. (2015). Negative regulation of ABA signaling by WRKY33 is critical for *Arabidopsis* immunity towards *Botrytis cinerea* 2100. *eLife* 4:e07295. doi: 10.7554/eLife.07295

Luo, D. L., Ba, L. J., Shan, W., Kuang, J. F., Lu, W. J., and Chen, J. Y. (2017). Involvement of WRKY transcription factors in abscisic-acid-induced cold tolerance of banana Fruit. *J. Agric. Food Chem.* 65, 3627–3635. doi: 10.1021/acs.jafc.7b00915

Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., and Zhang, S. (2011). Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexinbiosynthesis in *Arabidopsis*. *Plant Cell* 23, 1639–1653. doi: 10.2307/41433414

Née, G., Xiang, Y., and Soppe, W. J. (2017). The release of dormancy, a wake-up call for seeds to germinate. *Curr. Opin. Plant Biol.* 35, 8–14. doi: 10.1016/j.pbi.2016.09.002

Ouwerkerk, P. B., and Meijer, A. H. (2001). Yeast one-hybrid screening for DNA-protein interactions. *Curr. Protoc. Mol. Biol.* 55:12. doi: 10.1002/0471142727.mb1212s55

Parada, F., Noriega, X., Dantas, D. J., Bressansmith, R., and Perez, F. J. (2016). Differences in respiration between dormant and non-dormant buds suggest the involvement of ABA in the development of endodormancy in grapevines. *J. Plant Physiol.* 201, 71–78. doi: 10.1016/j.jplph.2016.07.007

Pérez, F. J., Rubio, S., and Ormeño-Núñez, J. (2007). Is erratic bud-break in grapevines grown in warm winter areas related to disturbance in mitochondrial respiratory capacity and oxidative metabolism? *Funct. Plant Biol.* 34, 624–632. doi: 10.1071/FP06272

Rubio, S., Noriega, X., and Perez, F. J. (2019). ABA promotes starch synthesis and storage metabolism in dormant grapevine buds. *J. Plant Physiol.* 234, 1–8. doi: 10.1016/j.jplph.2019.01.004

Rushton, D. L., Tripathi, P., Rabara, R. C., Lin, J., Ringler, P., Boken, A. K., et al. (2012). WRKY transcription factors: key components in abscisic acid signalling. *Plant Biotechnol. J.* 10, 2–11. doi: 10.1111/j.1467-7652.2011.00634.x

Rushton, P. J., Macdonald, H., Huttly, A. K., Lazarus, C. M., and Hooley, R. (1995). Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of α-Amy2 genes. *Plant Mol. Biol.* 29, 691–702. doi: 10.1007/BF00041160

Rushton, P. J., Somssich, I. E., Ringler, P., and Shen, Q. J. (2010). WRKY transcription factors. *Trends Plant Sci.* 15, 247–258. doi: 10.1016/j.tplants.2010.02.006

Seo, M., Hanada, A., Kuwahara, A., Endo, A., Okamoto, M., Yamauchi, Y., et al. (2006). Regulation of hormone metabolism in *Arabidopsis* seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. *Plant J.* 48, 354–366. doi: 10.1111/j.1365-313X.2006.02881.x

Singh, R. K., Miskolczi, P., Maurya, J. P., and Bhalerao, R. P. (2018). A tree ortholog of SHORT VEGETATIVE PHASE floral repressor mediates photoperiodic control of bud dormancy. *Curr. Biol.* 29, 1–6. doi: 10.1016/j.cub.2018.11.006

Singh, R. K., Svystun, T., AlDahmash, B., Jönsson, A. M., and Bhalerao, R. P. (2017). Photoperiod- and temperature-mediated control of phenology in trees-a molecular perspective. *New Phytol.* 213, 511–524. doi: 10.1111/nph.14346

Tylewicz, S., Petterle, A., Marttila, S., Miskolczi, P., Azeez, A., Singh, R. K., et al. (2018). Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* 360, 212–215. doi: 10.1126/science.aan8576

Vimont, N., Fouché, M., Campoy, J. A., Tong, M., Arkoun, M., Yvin, J. C., et al. (2019). From bud formation to flowering: transcriptomic state defines the cherry developmental phases of sweet cherry bud dormancy. *BMC Genomics* 12:974. doi: 10.1101/586651

Wang, F., Zhao, P., Zhang, L., Zhai, H., and Du, Y. (2019). Functional characterization of WRKY46 in grape and its putative role in the interaction between grape and phylloxera (*Daktulosphairavitifoliae*). *Hortic. Res.* 6:102. doi: 10.1038/s41438-019-0185-8

Wang, X., Guo, C., Peng, J., Li, C., Wan, F., Zhang, S., et al. (2019). ABRE-BINDING FACTORS play a role in the feedback regulation of ABA signaling by mediating rapid ABA induction of ABA co-receptor genes. *New Phytol.* 221, 341–355. doi: 10.1111/nph.15345

Weigel, D., and Glazebrook, J. (2002). *Arabidopsis*: a laboratory manual. *Eur. J. Soc. Theory* 10, 339–356. doi: 10.1177/1368431007080699

Wu, R., Wang, T., Warren, B., Allan, A. C., Macknight, R. C., and Varkonyi-Gasic, E. (2017). Kiwifruit SVP2 gene prevents premature budbreak during dormancy. *J. Exp. Bot.* 68, 1071–1082. doi: 10.1093/jxb/erx014

Xin, Z., and Browse, J. (2010). Cold comfort farm: the acclimation of plants to freezing temperatures. *Plant Cell Environ.* 23, 893–902. doi: 10.1046/j.1365-3040.2000.00611.x

Yang, Q., Niu, Q., Tang, Y., Ma, Y., Yan, X., Li, J., et al. (2019). PpyGAST1 is potentially involved in bud dormancy release by integrating the GA biosynthesis and ABA signaling in 'Suli' pear (*Pyruspyrifolia* White Pear Group). *Environ. Exp. Bot.* 162, 302–312. doi: 10.1016/j.envexpbot.2019.03.008

Yang, Q., Yang, B., Li, J., Wang, Y., Tao, R., Yang, F., et al. (2020). ABA-responsive ABRE-BINDING FACTOR3 activates DAM3 expression to promote bud dormancy in Asian pear. *Plant Cell Environ.* 43, 1360–1375. doi: 10.1111/pce.13744

Yordanov, Y. S., Cathleen, M., Strauss, S. H., and Busov, V. B. (2014). EARLY BUD-BREAK 1 (EBB1) is a regulator of release from seasonal dormancy in poplar trees. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10001–10006. doi: 10.1073/pnas.1405621111

Zheng, C., Acheampong, A. K., Shi, Z., Mugzech, A., Halaly-Basha, T., Shaya, F., et al. (2018a). Abscisic acid catabolism enhances dormancy release of grapevine buds. *Plant Cell Environ.* 41, 2490–2503. doi: 10.1111/pce.13371

Zheng, C., Halaly, T., Acheampong, A. K., Takebayashi, Y., Jikumaru, Y., Kamiya, Y., et al. (2015). Abscisic acid (ABA) regulates grape bud dormancy, and dormancy release stimuli may act through modification of ABA metabolism. *J. Exp. Bot.* 66, 1527–1542. doi: 10.1093/jxb/eru519

Zheng, C., Kwame Acheampong, A., Shi, Z., Halaly, T., Kamiya, Y., Ophir, R., et al. (2018b). Distinct gibberellin functions during and after grapevine bud dormancy release. *J. Exp. Bot.* 69, 1635–1648. doi: 10.1093/jxb/ery022

Zhou, C., Lin, Q., Lan, J., Zhang, T., Liu, X., Miao, R., et al. (2020). WRKY transcription factor OsWRKY29 represses seed dormancy in rice by weakening abscisicacid response. *Front. Plant Sci.* 11:691. doi: 10.3389/fpls.2020.00691

Zhu, Y., Liu, X., Gao, Y., Li, K., and Guo, W. (2021). Transcriptome-based identification of AP2/ERF family genes and their cold-regulated expression during the dormancy phase transition of Chinese cherry flower buds. *Sci. Hortic.* 275:109666. doi: 10.1016/j.scienta.2020.109666

Conflict of Interest: 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 17 August 2022
DOI 10.3389/fpls.2022.971846

OPEN ACCESS

EDITED BY
Yuexue Liu,
Shenyang Agricultural University, China

REVIEWED BY
Takeshi Kurokura,
Utsunomiya University, Japan
Chang Liu,
University of Hohenheim, Germany

*CORRESPONDENCE
Iraida Amaya
iraida.amaya@juntadeandalucia.es
Cristina Castillejo
cristina.c.mangado@juntadeandalucia.es

†These authors share senior authorship

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 17 June 2022
ACCEPTED 22 July 2022
PUBLISHED 17 August 2022

CITATION
Muñoz-Avila JC, Prieto C, Sánchez-Sevilla JF, Amaya I and Castillejo C (2022) Role of *FaSOC1* and *FaCO* in the seasonal control of reproductive and vegetative development in the perennial crop *Fragaria* × *ananassa*. *Front. Plant Sci.* 13:971846. doi: 10.3389/fpls.2022.971846

Role of *FaSOC1* and *FaCO* in the seasonal control of reproductive and vegetative development in the perennial crop *Fragaria* × *ananassa*

Julio C. Muñoz-Avila[1], Concepción Prieto[1], José F. Sánchez-Sevilla[1,2], Iraida Amaya [1,2*†] and Cristina Castillejo [1*†]

[1]Laboratorio de Mejora y Biotecnología, Instituto Andaluz de Investigación y Formación Agraria y Pesquera (IFAPA) Centro de Málaga, Málaga, Spain, [2]Unidad Asociada de I + D + i IFAPA-CSIC, Biotecnología y Mejora en Fresa, Málaga, Spain

The diploid woodland strawberry (*F. vesca*) represents an important model for the genus *Fragaria*. Significant advances in the understanding of the molecular mechanisms regulating seasonal alternance of flower induction and vegetative reproduction has been made in this species. However, this research area has received little attention on the cultivated octoploid strawberry (*F.* × *ananassa*) despite its enormous agronomical and economic importance. To advance in the characterization of this intricated molecular network, expression analysis of key flowering time genes was performed both in short and long days and in cultivars with seasonal and perpetual flowering. Analysis of overexpression of *FaCO* and *FaSOC1* in the seasonal flowering 'Camarosa' allowed functional validation of a number of responses already observed in *F. vesca* while uncovered differences related to the regulation of FaFTs expression and gibberellins (GAs) biosynthesis. While FvCO has been shown to promote flowering and inhibit runner development in the perpetual flowering H4 accession of *F. vesca*, our study showed that *FaCO* responds to LD photoperiods as in *F. vesca* but delayed flowering to some extent, possibly by induction of the strong *FaTFL1* repressor in crowns. A contrasting effect on runnering was observed in FaCO transgenic plants, some lines showing reduced runner number whereas in others runnering was slightly accelerated. We demonstrate that the role of the MADS-box transcription factor FaSOC1 as a strong repressor of flowering and promoter of vegetative growth is conserved in woodland and cultivated strawberry. Our study further indicates an important role of FaSOC1 in the photoperiodic repression of FLOWERING LOCUS T (FT) genes *FaFT2* and *FaFT3* while *FaTFL1* upregulation was less prominent than that observed in *F. vesca*. In our experimental conditions, FaSOC1 promotion of vegetative growth do not require induction of GA biosynthesis, despite GA biosynthesis genes showed a marked photoperiodic upregulation in response to long days, supporting GA

requirement for the promotion of vegetative growth. Our results also provided insights into additional factors, such as FaTEM, associated with the vegetative developmental phase that deserve further characterization in the future.

KEYWORDS

flowering, runnering, photoperiod, *Fragaria*, cultivated strawberry, *FaSOC1*, gibberellins, *FaCO*

Introduction

Strawberry (*Fragaria* spp.) belongs to the Rosaceae family and comprise different species among which, the commonly called woodland strawberry *F. vesca* (2n = 2× = 14) has become a genetic model for studying perennial development in Rosaceae (Hytönen and Kurokura, 2020). The cultivated strawberry, *Fragaria* × *ananassa* (2n = 8× = 56), is the most important berry crop in the world, with a global production over 8.8 M tones in 2020 (FAOSTAT 2022). Both species are closely related as an ancestor of *F. vesca* is the dominant progenitor of the octoploid strawberry (Edger et al., 2019). Breeding efforts in *F.* × *ananassa* are challenged by its octoploid nature and a general high heterozygosity, which result in many alleles contributing to trait variation and in the necessity of vegetative propagation to ensure the maintenance of superior genotypes.

Strawberries are perennial rosette plants that can be propagated both sexual (flowering) and asexually (vegetatively or clonally). Vegetative reproduction takes place through runners, also called stolons, which are elongated stems that grow horizontally above ground from which daughter plants arise. They are important for growers and breeders as they are used to propagate strawberry cultivars avoiding trait loss through recombination. However, runnering is also associated with a decrease in fruit yield, as it reduces the number of fruit-bearing shoots (Gaston et al., 2013; Heide et al., 2013; Tenreira et al., 2017). As a consequence, commercial strawberry fruit production is highly influenced by the balance between the two reproductive modes. One of the most important goals in strawberry breeding programs is increasing or maintaining high yield. A better understanding of the genetic and molecular factors that modulate the switch between inflorescence and runner development would facilitate the development of new cultivars with extended fruiting period and/or better adapted to local environments and changing climatic conditions, enabling fruit production under different photoperiods, temperatures or latitudes.

The flowering/runnering decision in strawberry depends on the fate of plant meristems. The stem or crown consists of short internodes produced from the shoot apical meristem (SAM). Each node harbors one leaf and an axillary meristem (AXM). Inflorescences are always formed terminally from the apical meristem (Darrow, 1966; Guttridge, 1985). After the emergence of the terminal inflorescence the crown vegetative extension continues from the uppermost lateral meristem, from which secondary and tertiary inflorescences could arise. On the other hand, a bud emerged from an AXM might remain dormant or activated to grow and develop either into a runner or into a new branch crown, which eventually can bear a terminal inflorescence (Heide et al., 2013; Perrotte et al., 2016b). The fate of the apical and axillary meristems is dictated by genetic and environmental conditions (Heide, 1977; Guttridge, 1985; Sønsteby and Nes, 1998; Bradford et al., 2010; Hytönen and Kurokura, 2020; Andrés et al., 2021).

Cultivated and woodland strawberries are classified based on their photoperiodic response into either seasonal (short-day; SF) or perpetual (day-neutral; PF) flowering types (Iwata et al., 2012; Gaston et al., 2013; Heide et al., 2013; Hytönen and Kurokura, 2020). During the fall, in response to short days (SD) and low temperatures, branch crowns from seasonal types emerge from AXMs and floral induction occurs at the SAMs of the main crown and older branch crowns. After adequate time in flowering-inducing conditions, plants become semi-dormant until sufficient winter chilling resumes vigorous growth. Then, terminal inflorescences emerge from previously induced meristems. Later on, under the long photoperiods (LD) and warm temperatures of summer, strawberry plants grow vegetatively and axillary buds differentiate into runners instead of branch crowns (Guttridge, 1958; Jonkers, 1965; Heide et al., 2013). Additionally, natural PF mutants in which flowering occurs all along the vegetative cycle have been identified in cultivated (*F.* × *ananassa*) and woodland (*F. vesca*) strawberries (Iwata et al., 2012; Gaston et al., 2013). Heide et al. (2013) studied flowering habit of PF types of woodland and cultivated strawberries in SD and LD conditions and concluded that they could be considered quantitative LD plants rather than day-neutrals.

The genetic control of photoperiodic flowering in both diploid woodland strawberry and octoploid cultivated strawberry have long been investigated (Brown and Wareing, 1965; Gaston et al., 2013). In *F. vesca*, genes affecting natural

variation on photoperiodic flowering and runnering have been identified. Two groups independently found recessive mutations in the *F. vesca* homolog of *TERMINAL FLOWER 1* (*FvTFL1*) as the underlying variation in the SF locus (SFL) on chromosome 6, causing perpetual flowering (Iwata et al., 2012; Koskela et al., 2012). The Runnering (R) locus on chromosome 2 causes runnerless plants and a deletion in the active site of the gibberellin (GA) biosynthetic gene *FvGA20ox4* was identified in all *r* mutants (Tenreira et al., 2017). A later study confirmed that FvGA20ox4 is indispensable for runner development and under tight environmental regulation (Andrés et al., 2021). In cultivated strawberry, the PF trait has been shown to be controlled by the major quantitative trait locus (QTL) *FaPFRU,* which has been mapped to chromosome 4B (Gaston et al., 2013; Castro et al., 2015; Perrotte et al., 2016a). The underlying gene for *FaPFRU*, different from *TFL1*, is still unknown and should act as a dominant and positive regulator of flowering (Perrotte et al., 2016a). However, the strong floral repressor role of TFL1 was confirmed in the cultivated strawberry, as silencing of *FaTFL1* in a SF cultivar was sufficient to induce PF (Koskela et al., 2016). Although the *FaPFRU* QTL also displays important negative effects on runnering (Perrotte et al., 2016a), the genetic basis of runnering in cultivated strawberry appears to be complex, as weak additive genetic effects and many small effect QTLs have been detected (Simpson and Sharp, 1988; Hossain et al., 2019).

As in other species, besides photoperiod, temperature has an important effect on flowering initiation in strawberry (Heide, 1977; Sønsteby and Heide, 2006; Heide et al., 2013; Rantanen et al., 2015). In both SF and PF genotypes, flowering can be promoted at temperatures below 10–13°C, behaving essentially as day-neutral, while at temperatures above 23–25°C flowering can be inhibited even under inductive photoperiods. In *F. vesca*, this temperature-dependent induction or repression of flowering relies on FvTFL1, which expression is regulated by ambient temperature, while in mild conditions, *FvTFL1* is under photoperiodic control (Rantanen et al., 2015; Koskela et al., 2017; Hytönen and Kurokura, 2020).

The model plant *Arabidopsis* perceives the increase in day length by a complex mechanism that involves the transcriptional and post-transcriptional regulation of the B-Box transcription factor CONSTANS (CO), which eventually triggers the expression of *FT* at the correct season. FT induction occurs in the leaf companion cells and it is transported through the phloem to the apical meristem, changing the fate of the tissue (Suárez-López et al., 2001; Valverde et al., 2004; Valverde, 2011). This CO–FT module works in most species analyzed, however, the final effect of this signal diverges in different species. For example, in rice, a SD plant, the CO-FT module functions as an activator in SD but a repressor in non-inductive LDs (Hayama et al., 2003). By contrast, in legumes such as *Medicago* or pea, CO-like genes may not have any role in the integration of flowering in response to photoperiodic cues (Serrano-Bueno et al., 2017).

Most of what we know about the molecular events that regulate flowering and runnering in strawberry comes from studies using the diploid model *F. vesca,* in which the LD-activated FvFT1–FvSOC1–FvTFL1 module has a pivotal role in the repression of flowering, in a *FvTFL1* wild type (WT) background, or activation of flowering, in a *fvtfl1* mutant background (Hytönen and Kurokura, 2020). FvCO protein has been suggested as part of the photoperiod measurement system necessary for *FvFT1* induction. In the PF *F. vesca* accession Hawaii-4 (H4), which lacks a functional allele of the floral repressor *FvTFL1*, overexpression of *FvCO* leads to *FvFT1* induction in leaf that correlates with upregulation of SUPPRESSOR OF THE OVEREXPRESSION OF CONSTANS1 (*FvSOC1*) and the floral meristem identity gene APETALA1 (*FvAP1*) at the SAMs, and eventually promotes flowering under LD, whereas silencing of *FvCO* has the opposite effect (Koskela et al., 2017). In SF genotypes, upregulation of *FvTFL1* by FvSOC1 prevents flower induction under LD conditions (Mouhu et al., 2013). In addition, *FvSOC1* promotes vegetative growth and runnering by activating the expression of several gibberellin (GA) biosynthesis genes, including *FvGA20ox4*.

Three *FT* genes have been detected in woodland and cultivated strawberry. In *F. vesca*, induction of *FvFT1* in leaves is regulated by light quality in addition to photoperiod (Koskela et al., 2012; Rantanen et al., 2014; Prisca et al., 2022). Leaf-expressed FvFT2 has been shown to act as a mobile signal for flowering under SD photoperiod, while *FvFT3* was not detected in leaf and may promote plant branching, thus increasing flower number and yield in *F. vesca* (Gaston et al., 2021). Expression analyses of *FaTFL1*, *FaFT1*, *FaFT2,* and *FaFT3* in *F.* × *ananassa* SF cultivars grown under different photoperiods and temperatures suggest that *FaFT1/FaTFL1* and *FaFT3* play important roles on the environmental repression and induction of flowering, respectively (Nakano et al., 2015; Koembuoy et al., 2020).

The tradeoff between vegetative propagation and flowering in both cultivated and woodland strawberries is also regulated by gibberellins (GAs). Exogenous GA application to SD-grown strawberry can mimic the effect of LD conditions, promoting runner initiation and inhibiting flowering (Thompson and Guttridge, 1959; Nishizawa, 1993; Black, 2004). On the other hand, treatment with the GA biosynthesis inhibitor prohexadione-calcium (Pro-Ca) enhances branch crown at the expense of runner formation (Nishizawa, 1993; Black, 2004; Hytönen et al., 2009). These observations were genetically supported in *F. vesca* with the identification of a 9-bp deletion in the GA20-oxidase gene *FvGA20ox4* as the runnerless (r) mutation (Tenreira et al., 2017). The importance of GA signaling for runner production in *F. vesca* was further demonstrated

when the DELLA protein FvRGA1, a negative regulator of GA signaling, was identified as a key protein controlling runner formation and the causal loci behind the suppressor of runnerless (srl) mutation (Caruana et al., 2018; Li et al., 2018).

As aforementioned, genetic analysis of the PF trait in *F. vesca* and *F.* × *ananassa* led to the identification of different loci in each species, suggesting that differences in flowering time regulation may exist between them. At the same time, the function of factors like the flowering repressor TFL1 is conserved between the diploid model *F. vesca* and *F.* × *ananassa* (Koskela et al., 2016). In order to establish how much of the knowledge acquired in *F. vesca* can be translated to the commercially cultivated species, the role of other proteins such as FaCO and FaSOC1, as well as additional regulation of flowering, need to be further investigated in the octoploid. It is still unclear whether FT2 or FT3 act as mobile or SAM-based floral promoters under SD conditions in cultivated strawberry. In this work, using a combination of expression analysis of selected flowering time genes and functional validation by transgenesis, we have been able to confirm some genetic responses to photoperiod already described in *F. vesca* but also to uncover differences in the mechanisms regulating the transition to flowering, particularly in aspects related to GA metabolism and signaling.

Materials and methods

Plant material and growth conditions

All *F.* × *ananassa* plants used in this study were maintained and grown in shaded greenhouses (standard or for GMOs in the case of transgenic plants) under natural sunlight and temperature conditions at IFAPA, Málaga, Spain. Temperature conditions were recorded in an outdoor meteorological station 1500 m away from the shaded greenhouses. For the different experiments, plants were clonally propagated from runners during June-September and potted in October/November in 22 cm pots with a mixture of universal substrate and river sand (3:1 v/v).

For analysis of gene expression in the different plant tissues, leaf, root, crown, flower and green, white and ripe fruit were sampled as shown in **Supplementary Figure 1**. They were sampled under SD conditions (9 h38 min/14 h21 min day/night; average maximum and minimum temperature from previous 4 weeks: 17.7/9.7°C) from cultivar Chandler. Three biological replicates were collected, each consisting of 4–6 plants.

For comparison of short day and long day (SD/LD) conditions, leaf and crown samples from seasonal flowering (SF) cultivar Chandler and perpetual flowering (PF) cultivar Selva were harvested in December 21st 2012 (9 h38 min/14 h21 min day/night; average maximum and minimum temperature from previous 4 weeks: 17.7/9.7°C) and June 21st 2013 (14 h40 min/9 h19 min day/night; average maximum and minimum temperature from previous 4 weeks: 26.7/16.5°C) for SD and LD photoperiods, respectively. Three biological replicates were collected at Zeitgeber time (ZT) 6, each consisting of 4–6 plants.

For analysis of circadian rhythmic expression, leaf samples from 'Chandler' and 'Selva' were collected in two separate experiments, at SD and LD conditions. SD sampling took place on December 22nd 2013 (9 h38 min/14 h21 min day/night; average maximum and minimum temperature from previous 4 weeks: 18.2/8.8°C) and LD on June 23rd 2014 (14 h40 min/9 h19 min day/night; average maximum and minimum temperature from previous 4 weeks: 29.6/19.4°C). The first time point was sampled at dawn and then at 4, 8, 12, 16, 20, and 24 h. Three biological replicates consisting of leaves from at least three plants were collected at each time point.

For evaluation of transgenic phenotypes, 9 clones of each transgenic line were propagated from runners as described above. Nine clones from two pGUS lines (pGUS2 and pGUS3) and from Camarosa wild type (WT) were grown and used as controls. For gene expression studies in 35S:*FaSOC1* lines, young leaf (first leaf with fully expanded leaflets) and crown tissues from 3 biological replicates were collected under natural SD conditions (November 29th 2017, 9 h52 min/14 h17 min day/night; average maximum and minimum temperature from previous 4 weeks: 21.1/10.6°C) at ZT 3, each replicate being a pool from 3 plants. As initial evaluation of T_0 transgenic plants revealed similar phenotypes in lines pGUS2, pGUS3, and WT non-transformed plants, only the pGUS3 control was kept as control for gene expression studies.

Phenotypic data were analyzed using one-way ANOVA followed by Tukey's test. A value of $P < 0.05$ was considered as statistically significant. Graphs and statistical analyses were performed using the GraphPad Prism 8 software.

Gene expression analysis

Total RNA was extracted from three biological replicates following a CTAB protocol (Gambino et al., 2008) from 200 mg of frozen powdered samples from vegetative tissues or 300 mg from fruit samples. Residual DNA from RNA samples was removed using Invitrogen™ TURBO DNA-free™ Kit and cDNA was synthesized from 1 μg of RNA using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). Gene expression levels were analyzed in triplicate by quantitative real-time PCR (qPCR) with the SsoFast EvaGreen supermix (Bio-Rad) and a CFX96 real-time PCR system (Bio-Rad) using a standard two-step program of 40 cycles, annealing at

60°C. The relative expression was calculated using the geometric mean of the housekeeping genes DBP and GAPDH for normalization and the $2^{-\Delta\Delta Ct}$ method (Pfaffl, 2001) unless stated otherwise in the figure legend. Primer sequence information for all analyzed genes is listed in **Supplementary Table 1**.

GraphPad Prism was used for statistical analysis. If data points passed the D'Agostino-Pearson omnibus K2 normality test ($P < 0.001$) an ordinary one-way ANOVA was employed followed by either a Fisher's LSD test to compare preselected pairs of columns (SD/LD genotypes and GA treatment experiments), or a Tuckey's test to compare each column mean with every other mean (tissue expression) or, when comparing each group to a control group, a Dunnett's test (expression in transgenic lines). When data points didn't pass the normality test, *p*-values were determined by a Kruskal Wallis test, followed by Dunn's post-test for multiple group comparisons. Mean and SEM of the 3 biological replicates were plotted. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

Plasmid construction

For overexpression of *FaCO* and *FaSOC1*, ORFs were cloned from cv. Selva leaf and crown tissues, respectively. PCR amplification was performed using Pfu polymerase 5-prime and DNA fragments cloned in pGEMT-easy (Promega) and sequenced. To generate the *35S:CO* and *35S:FaSOC1* overexpression constructs, each gene was cloned into the *Sal*I and *Xba*I restriction sites of the pBINPLUS vector (van Engelen et al., 1995). A *35S:GUS* construct in the same pBINPLUS vector was used as control.

Transgenic plants

Leaf explants from *in vitro* 'Camarosa' plants were transformed with *Agrobacterium tumefaciens* LBA4404 harboring the *35S:FaCO, 35S:FaSOC1* or *35S:GUS* constructs as previously described (El Mansouri et al., 1996). Transgenic shoots were selected on medium containing 50 mg/L kanamycin and 500 mg/L carbenicillin and grown *in vitro* in a culture chamber under cool-white light (at 15 mE) and a long-day photoperiod (16-h light/8-h dark) at 22°C. Resistant plantlets were acclimated to soil conditions, then transferred into 22 cm pots and grown in confined greenhouse under natural environmental conditions. Preliminary evaluations on the first year were performed with only one replicate of each transgenic line. For successive years, plants were propagated during the summer and 9 plants of each line and controls were transferred to the greenhouse. Phenotype evaluation was performed during the growing season, from December to September.

Gibberellin treatment

Gibberellin treatment was performed on 6 weeks old clonally propagated *F.* × *ananassa* 'Camarosa' plants grown under natural greenhouse conditions. Six plants were used for each of the 3 biological replicates. GA_3 (SIGMA S7645) was first dissolved in ethanol at 50 mg/mL and a diluted 1:1000 working solution was made in 0.1% Tween-20 (GA_3 final concentration 50 mg/L). Controls were treated with a 0.1% Tween-20 0.1% ethanol solution. Plants were sprayed to runoff twice, with the first application conducted at day 0 and a second one at day 2. Crown and young leaf samples were collected at day 3 (October 7th, 11 h41 min/13 h18 min day/night, average maximum and minimum temperature from previous 4 weeks: 26.4/16.4°C) at ZT 3.

Results

Photoperiodic regulation of flowering time genes in seasonal flowering and perpetual flowering *F.* × *ananassa* cultivars

As a starting point for our study on the genetic pathways controlling the transition to flowering in response to day length in cultivated strawberry, we compared the expression of selected flowering time genes in the SF and PF cultivars Chandler and Selva, respectively. Light and photoperiod sensing occurs mainly in leaves and this signal is transmitted to plant meristems to determine their fate (Song et al., 2015). The strawberry plant shoot, also called crown, contains the SAM at terminal position and AXMs at the basis of each leaf. Expression analyses were performed in these two tissues, leaves and crowns (**Supplementary Figure 1**), collected under natural SD or LD photoperiods. The phenological phase of the meristems was determined by evaluating the expression of the major floral meristem identity genes *FaLFYa, FaAP1,* and *FaFUL,* as well as the repressor of floral transition *FaTFL1* (**Figure 1A**). Expression levels of *FaLFYa* and *FaAP1* were higher in crowns from plants collected under the inductive SD photoperiods in both cultivars compared to LD, whereas *FaFUL* induction was only detected in crowns from 'Selva.' On the other hand, the floral inhibitor *FaTFL1* was downregulated in SD crowns, coinciding with *FaAP1* and *FaLFY* induction. However, *FaTFL1* upregulation under LD was only observed in crowns from 'Selva.' Rather than cultivar differences in the signaling pathways, these results are more likely to reveal

differences in the timing of phase transitions between them. Altogether, these results suggest that only plants grown in SD presented floral meristems in their crowns, and that the transition to flowering, and later in LD to the vegetative phase, lagged behind in 'Chandler' compared to 'Selva.' Additionally, the floral meristem identity genes were not upregulated in the LD sampled meristems from the PF cultivar 'Selva,' indicating that the transition to the second reproductive phase had not taken place when samples were collected in late June.

In leaf samples harvested under LD, coinciding with the beginning of summer, we observed a clear induction of *FaCO* in both cultivars that correlates with higher *FaFT1* expression in leaves (**Figure 1B**). In 'Chandler,' *FaCO* upregulation in leaves parallels a drop in *FaTEM* expression in the same tissue (**Figure 1B**). Although not statistically significant, a similar trend in *FaTEM* expression was observed in 'Selva.' As proposed in *F. vesca* (Koskela et al., 2012; Rantanen et al., 2014), in cultivated strawberry FaFT1 induced in leaves might act as a long-distance signal and move to the apical meristems. However, in the presence of functional FaTFL1, the overall effect is floral repression and promotion of the vegetative state, as reflected by the low transcript level of the floral markers *FaLFYa* and *FaAP1* (**Figure 1A**). Noteworthy, the most prominent change in gene expression under LD in the two cultivars was the strong *FaSOC1* induction detected both in leaves and crowns (**Figure 1B**).

In the transition to winter, under SD, *FaFT2* and *FaFT3* expression showed a sharp peak in 'Chandler' leaves. In 'Selva,' a moderate similar trend was detected, suggesting the rise in expression of these genes might be transient and could have occurred earlier in 'Selva' (**Figure 1B**). Additionally, under flowering inductive SD, all three *FaFT* genes were induced in crowns of both cultivars, although *FaFT2* induction was only clearly detected in 'Chandler' (**Figure 1B**). *FaFTs* upregulation in the meristems is compatible with them acting upstream of the floral identity genes *FaLFYa* and *FaAP1* and thus promoting floral induction and inflorescence development under SD in cultivated strawberry (**Figure 1A**).

FaCO and FaSOC1 spatial and diurnal expression in cultivated strawberry

In the two *F.* × *ananassa* cultivars analyzed in this study, *FaCO* and *FaSOC1* expression is photoperiodically regulated. LD induced the expression of both genes, suggesting *FaCO* and *FaSOC1* might be involved in the regulation of the seasonal alternation between flowering and runnering in response to day-length. Therefore, we decided to focus our work on these two transcription factors.

First, we analyzed the expression of *FaCO* and *FaSOC1* in different tissues collected under SD from cultivar Chandler. As shown in **Figure 2A**, *FaCO* expression was specific to the aerial tissues of the plant, with leaves showing the highest *FaCO* transcript levels. This high expression is compatible with FaCO forming part of the photoperiod measurement system in leaves, as described in other species. Elevated *FaCO* expression was also detected in flowers and decreased during fruit development and ripening. In the case of *FaSOC1,* transcripts were detected in all tissues analyzed, although expression was highest in leaf and root, followed by crown, and it was markedly lower in reproductive organs (**Figure 2B**).

Then, the circadian expression of *FaCO* and *FaSOC1* was analyzed in leaves collected from the SF and PF accessions 'Chandler' and 'Selva' grown under SD and LD photoperiods (**Figures 2C–F**). *FaCO* daytime expression rhythm was similar to the one described for its ortholog *FvCO* (Rantanen et al., 2014), showing a peak of expression at the end of the dark period then quickly declining at dawn. The same pattern was observed under LD and SD photoperiods and in the two genotypes analyzed, and thus, the different flowering habits cannot be explained by differences in *FaCO* diurnal pattern of expression.

On the contrary, diurnal *FaSOC1* expression did not show any clear rhythm in none of the photoperiods tested nor cultivar-dependent differences. These results stand in contrast to previous observations by Koskela et al. (2016) in the SD cultivar Honeoye. They did not observe any diurnal oscillation under SD either, but in LD *FaSOC1* was slightly up-regulated in the morning. As occurred in the case of *FaCO*, the differences in flowering behavior between genotypes can't be explained by different circadian expression patterns of *FaSOC1*.

Effect of constitutive overexpression of *FaCO* on flowering and runnering in *F.* × *ananassa*

To further investigate the molecular mechanisms involved in the tradeoff between runnering and flowering and the role of *FaCO* in these processes in cultivated strawberry, we generated transgenic plants of the SF cultivar Camarosa ectopically expressing *FaCO* under the control of the CaMV 35S promoter. *FaCO* cDNA was amplified from leaf samples and sequence analysis revealed it corresponds to homoeolog *FxaC_23g53510* from 'Camarosa' chromosome Fvb6-2 (Edger et al., 2019). Several putative transgenic shoots from independent transformation events were obtained, from which nine PCR positive lines were acclimated and transferred to the greenhouse for initial evaluation and selection. *FaCO* overexpression was confirmed by RT-qPCR, with transgenic lines showing expression levels 7–80 times higher than the control (**Supplementary Figure 2A**). Flowering initiation was evaluated in these plantlets coming directly from tissue culture and therefore only one clone per transgenic line was available. A wide variation in the number of days before flowering

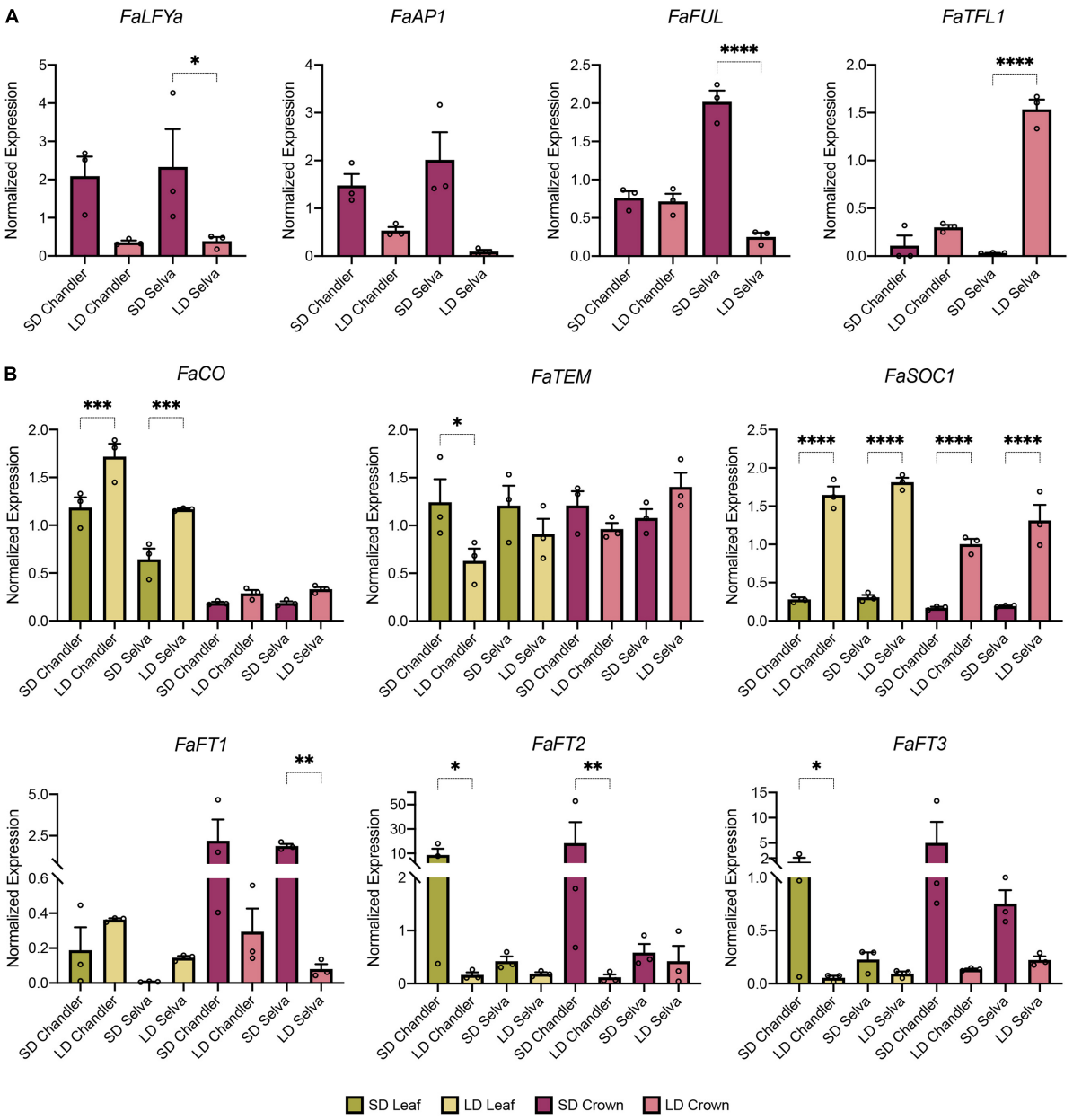


FIGURE 1
Transcriptional response of flowering time genes to short (SD) and long (LD) days in the seasonal flowering (SF) cultivar Chandler and the perpetual flowering (PF) cultivar Selva. **(A)** Expression of floral meristem identity genes and *FaTFL1* in crowns. **(B)** Expression of flowering time genes in leaves (green) and crowns (purple). Expression was analyzed by qRT-PCR. Graphs show the average of normalized values from three biological replicates ± SEM. In each bar, dots represent the normalized value of each biological replicate. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

was found among the nine transgenic lines (Supplementary Figure 2B). With the exception of line CO OverExpressor (COE)-14, all transgenic lines were characterized by less vigor, showing smaller plant size and slightly smaller leaves. Runner capacity was also affected at this stage. Three of the nine transgenic lines evaluated (lines COE-22, -40, and -42) were not able to produce runners, three had an intermediate phenotype (COE-12, -15, and -21) and only three (COE-11, -13, and -14) generated a similar number of runners than the controls, although the severity of this phenotype did not correlate with transgene expression level (Supplementary Figure 2A).

Lines unable to runner cannot be propagated and therefore are lost, hampering further evaluation of the effect of

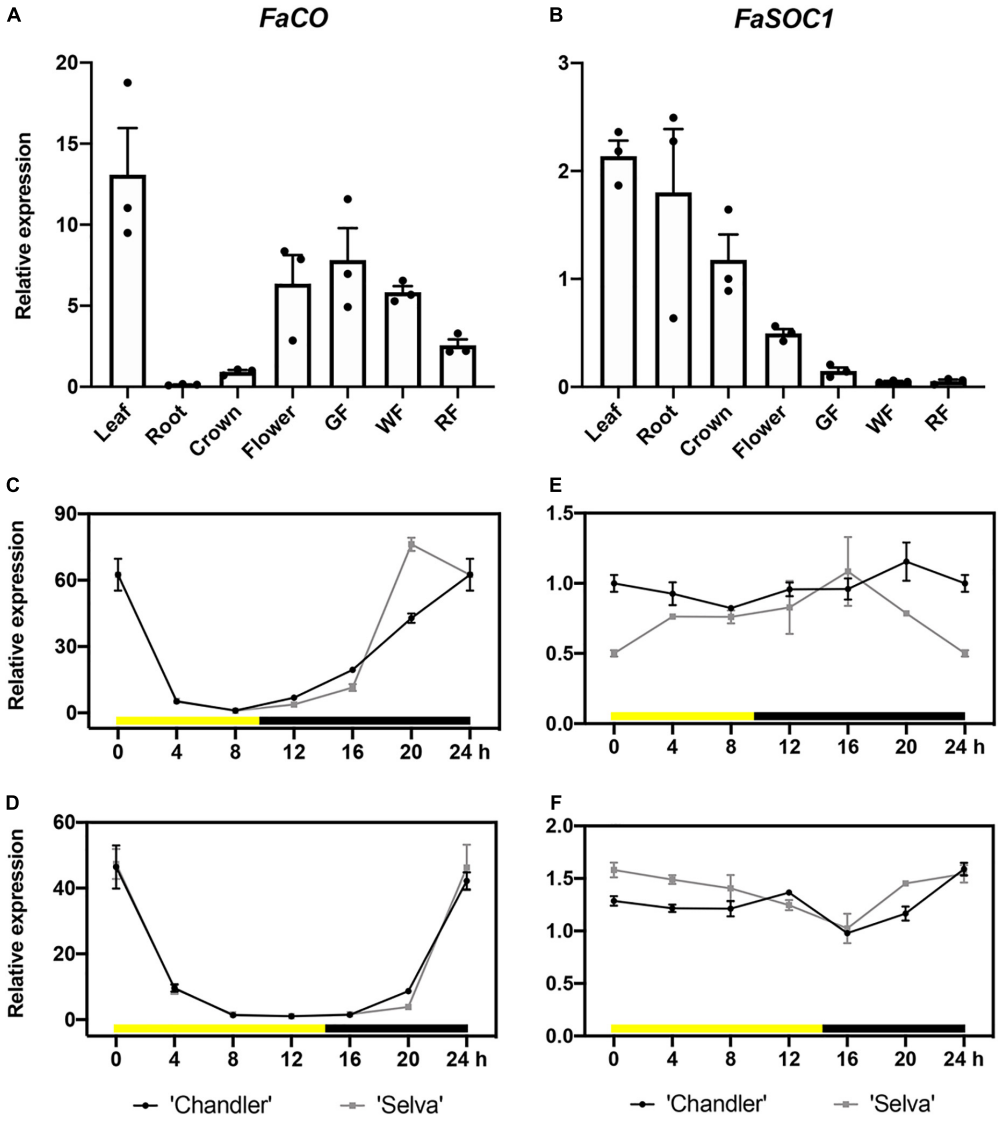


FIGURE 2
Spatial and diurnal expression of *FaCO* and *FaSOC1* by RT-qPCR. Relative expression levels of *FaCO* **(A)** and *FaSOC1* **(B)** in different tissues or organs. mRNA diurnal rhythms of *FaCO* under SD **(C)** and LD **(D)**. mRNA diurnal rhythms of *FaSOC1* under SD **(E)** and LD **(F)**. The first sample was taken at dawn: 0 h Zeitgeber time (ZT). Relative gene expression levels were normalized to the expression of *FaGAPDH* and crown tissue **(A,B)**, 8 h **(C,D)**, or 16 h **(E,F)** were used as reference samples. Data are means ± SEM of three biological replicates.

FaCO overexpression on runnering capacity beyond the first generation. Among the lines able to runner, COE-11, -12, and -13 were chosen as representative of the phenotypic variability observed in T0 lines and vegetatively propagated for further analyses. A total of nine plantlets per transgenic or control lines (WT and 35S:*GUS*) were grown in a greenhouse under natural light and temperature conditions. The three selected lines maintained a slightly smaller plant size throughout next generations (Figure 3). Winter flower initiation was slightly delayed in COE lines compared to control plants (Figure 3A), although statistically significant differences were only observed between pGUS control and lines COE-12 and 13, which flowered an average of 16 or 14 days later, respectively. Despite this subtle delay in the emergence of the first flowers, no differences in the cumulative number of flowers were observed (Figure 3B). Following the flowering season, runner initiation occurred slightly sooner in lines COE-11 and 13 (Figure 3C), but both of them ended up producing a similar number of runners. On the contrary, line COE-12, that initiated runners at the same time than controls, produced significantly less number (Figure 3D), as already observed during the first season (Supplementary Figure 2A). Altogether, these results point to a marginal and contrasting role of FaCO in the

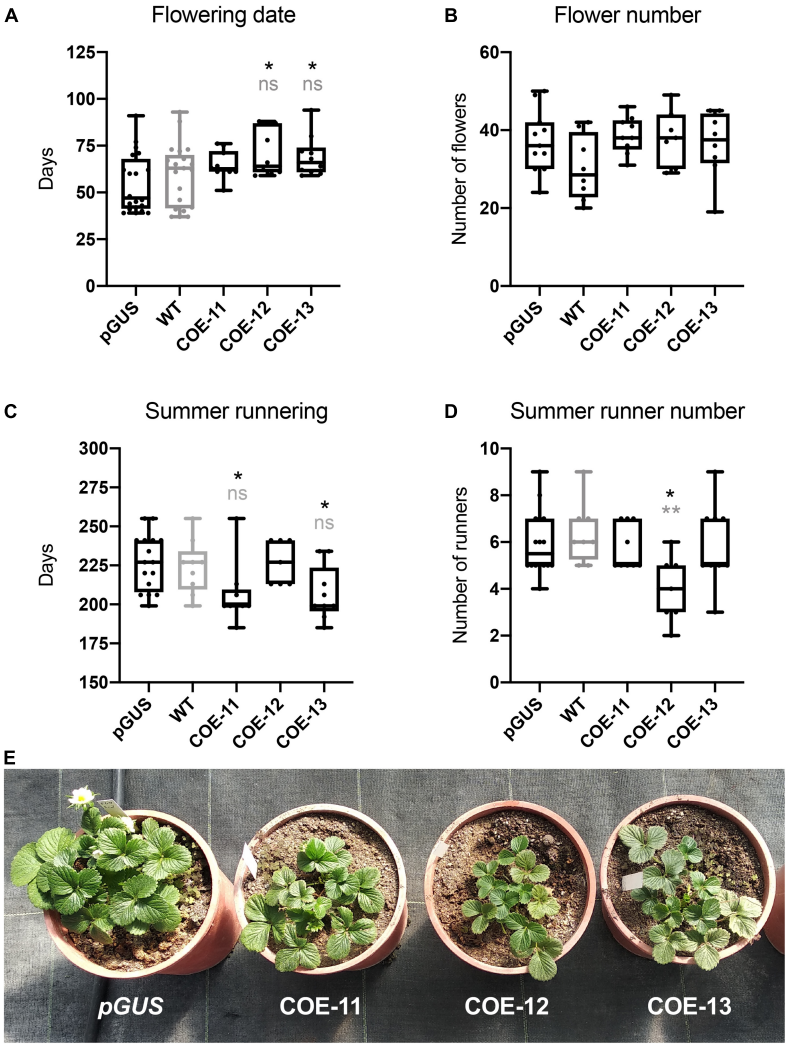


FIGURE 3
Phenotype of 35S:*FaCO* lines (COE) in comparison to control plants. **(A)** Number of days to flowering from November 20th. **(B)** Total flower number (from November 20th until September 1st). **(C)** Number of days until the first summer stolon. **(D)** Number of runners per plant. **(E)** Phenotype of 35S:*FaCO* lines in comparison to control GUS plants. One representative replicates is shown. Boxes span the 25th and 75th percentiles and the middle line represents the median. Whiskers (T-bars) are the minimum and maximum values. Asterisks indicate significant differences between pGUS (in black) or 'Camarosa' wild type (WT; in gray) and 35S:*FaCO* overexpressor plants; statistical significance was determined by ANOVA and Tukey's test (*$P < 0.05$, **$P < 0.01$).

control of meristem fate, at least under the growing conditions used in this work.

Overexpression of *FaSOC1* delays flowering and promotes vegetative development in *F.* × *ananassa*

FvSOC1, the *FaSOC1* ortholog from the diploid relative *F. vesca*, is known to alter the balance between runnering and flowering toward runnering, acting as a central hub where the fate of the meristem is established (Mouhu et al., 2013). To determine whether *FaSOC1* has a similar role in the cultivated strawberry, it was overexpressed in *F.* × *ananassa* cv. Camarosa. *FaSOC1* cDNA used for transformation was amplified from crowns and corresponded to the homoeolog *FxaC_25g18220* from 'Camarosa' chromosome Fvb7-2 (Edger et al., 2019). Eight PCR positive lines were selected and transferred to soil, all of them showing markedly higher *FaSOC1* expression in leaves than the control (**Supplementary Figure 3A**).

In a first-year evaluation, all 35S:*FaSOC1* lines exhibited increased vegetative development except line SOC1 OverExpressor (SOE)-1, the one with lower *FaSOC1* expression (**Supplementary Figure 3C**). A correlation between transgene expression level and phenotypic severity was observed and thus, lines SOE-20 and -21, the higher overexpressors, also showed

the most severe phenotype. As occurred in 35S:*FaCO* T_0 plants, a wide variability in flowering time was found among the eight 35S:*FaSOC1* lines. Four of them were late flowering, with SOE-7, -9, and -20 being particularly late, as they flowered at least 60 days later than the controls (**Supplementary Figure 3B**).

Lines SOE-7, -9, -20, and -21 were clonally propagated from runner cuttings. Nine plants per line, except SOE-7, were subjected to further molecular and phenotypic evaluation during the growing season, from December to September, in successive years. The number of days until the emergence of the first flower and the total number of flowers per plant were scored, showing *FaSOC1* overexpression had an evident impact on flowering induction (**Figure 4**). All 35S:*FaSOC1* lines flowered later than the controls (**Figure 4A**). Lines SOE-9, -20, and -21 flowered an average of 20, 25 and 9 days later than the WT, which in turn flowered 4 days later than the pGUS control, although statistically significant differences were only observed between both controls and SOE-9 and -20. Additionally, SOE-20 presented a significantly reduced number of flowers compared to control lines.

A very distinctive phenotype of the four SOE-7, -9, -20, and -21 lines was the production of runners in winter, immediately after separation from mother plants and before floral induction (**Figure 4H**). During summer LD, the three evaluated lines also produced runners significantly earlier than the controls (**Figure 4C**). Control lines started runners 223 or 227 days after planting while transgenic lines produced runners after 188–203 days. Although transgenic lines generated a slightly higher number of runners per plant, no significant differences were observed. The more vigorous appearance of 35S:*FaSOC1* lines was not a consequence of higher amount or larger leaves but rather more erect and taller plants. Although all of them were higher than the controls (**Figure 4H**), statistically significant differences were only observed in SOE-9 and -20 (**Figure 4E**). This phenotype was the reflection of considerably longer leaf petioles displayed by all SOE lines (**Figure 4F**), which also presented longer inflorescence peduncles (**Figure 4G**).

Ectopic and constitutive overexpression of *FaSOC1* negatively affected floral and fruit development (**Supplementary Figure 4**), increasing the incidence of malformed fruits and aborted flowers. Flowers within the inflorescences were more compact and presented defects in the sepal and petal whorls. Sepals were bigger whereas petals, although present, were reduced in size and senesced prematurely compared to the WT. Fruits were in general smaller and in the majority of cases did not develop into fleshy fruits. Line SOE-20, the one with higher transgene overexpression and stronger vegetative phenotype, was also the most affected in fruit development. The majority of SOE-20 fruits got completely dry in the green stage and only a small percentage developed further into dry pinkish fruits that remained attached to the plant. Line SOE-9 presented an intermediate phenotype and developed few wild type ripe fruits, while about half the fruits in line SOE-21 ripened into red fleshy fruits. Noticeably, achenes remained green in all three transgenic lines (**Supplementary Figure 4**).

Taken together, the photoperiodic pattern of *FaSOC1* expression and the phenotype of 35S:*FaSOC1* lines, point to *FaSOC1* acting as a negative regulator of flower induction and promoter of vegetative growth in cultivated strawberry.

Flowering-related genes were differentially expressed in 35S:*FaSOC1* lines

Since the transgenic plants overexpressing *FaSOC1* have a late flowering phenotype, the expression of flowering time genes, including *FaCO*, *FaTEM*, *FaFT1-3*, *FaTFL1*, *FaLFYa*, *FaAP1,* and *FaFUL*, was tested in all of them. Expression analysis were performed in young leaf and crown tissue collected under natural SD conditions. Overexpression of *FaSOC1* in leaf and crown of transgenic plants used in these experiments was confirmed by RT-qPCR (**Figure 5A**). As observed in the first season (**Supplementary Figure 3A**), SOE-20 was the line with highest transgene expression, followed by SOE-21 and then SOE-9.

FaCO expression levels were similar in 35S:*FaSOC1* and control plants in the two tissues analyzed (**Figures 5B,C**), supporting the model proposed in *F. vesca* that places *FvCO* upstream of *FvSOC1* (Kurokura et al., 2017). However, *FaTEM* was downregulated in leaves of all three SOE lines, although differences were only statistically significant in SOE-21. *FaTEM* expression in crowns, whereas not as evident as in leaves, also tended to be downregulated in the three transgenic lines (**Figures 5B,C**). The only *FaFT* family member detected in leaves of these SD grown plants was *FaFT1,* but its expression level was low and not affected by *FaSOC1* overexpression (**Figure 5B**).

The expression of the floral meristem genes *FaLFYa*, *FaAP1,* and *FaFUL* was significantly downregulated in *35S:FaSOC1* crowns compared to the control (**Figure 5C**), in agreement with their respective phenological stage. Whereas controls were flowering, SOE lines showed active vegetative development at the time of sampling. Additionally, *FaTFL1* expression was induced in lines SOE-20 and -21 (**Figure 5C**), supporting *FaTFL1* activation downstream of *FaSOC1* also happens in *F.* × *ananassa* as previously described in *F. vesca* (Mouhu et al., 2013).

Along with expression changes in floral meristem identity genes, the most dramatic changes in gene expression observed in 35S:*FaSOC1* crowns affected *FaFT2* and *FaFT3*, with transcript levels of both genes greatly reduced in all SOE lines (**Figure 5C**).

In contrast, *FaFT1* expression was similar in SOE and control lines (**Figures 5B,C**).

FaSOC1 overexpression does not activate gibberellin biosynthesis to promote vegetative growth

'Camarosa' plants overexpressing *FaSOC1* showed a characteristic elongated phenotype and initiated runners during winter short days, as previously observed in *F. vesca* overexpressing *FvSOC1* (Mouhu et al., 2013), resembling the effect of GA-treated plants (Thompson and Guttridge, 1959; Hytönen et al., 2008). Indeed, in *F. vesca*, FvSOC1 positively regulates the expression of the GA biosynthetic genes *FvGA20ox* and *FvGA3ox* required for runner differentiation (Mouhu et al., 2013). In order to evaluate if *FaSOC1* modulates GA metabolism, the expression of *FaGA20ox2, FaGA20ox4, FaGA3ox1, FaGA3ox2,* and *FaGA2ox1* was analyzed in leaves and crowns of SOE lines grown under SD.

FaGA20ox2 and *FaGA3ox2* were not detected in any of the tissues analyzed, neither in control nor in transgenic lines. *FaGA20ox4* was not detected in leaves but, although close to detection limits, it could be amplified from crowns. In lines SOE-9 and SOE-21 *FaGA20ox4* transcript level was similar to the control (**Figures 6A,B**). Only in the most severe SOE-20, *FaGA20ox4* tended to be upregulated but transcript levels were highly variable among the three biological replicates and thus not statistically significant (**Figure 6B**). *FaGA3ox1* expression in leaves and crowns was similar in SOE and control lines (**Figures 6A,B**).

As active GA levels are the result of their biosynthesis as well as their deactivation rates, *FaGA2ox1* expression, an oxidase involved in GA deactivation, was also analyzed. *FaGA2ox1* transcripts were detected in leaves and crowns but the expression level, as happened with *FaGA20ox4 and FaGA3ox1,* was not affected by *FaSOC1* overexpression (**Figures 6A,B**). Hence, our results do not support a role for FaSOC1 as an inducer of active GA accumulation in *F.* × *ananassa*, as was shown for its ortholog FvSOC1 in *F. vesca* (Mouhu et al., 2013).

Alternatively, as a similar constitutive GA response phenotype was described in the suppressor of runnerless (srl) *F. vesca* mutant (Caruana et al., 2018), the vegetative phenotype of SOE lines could also be explained by lower expression of the GA signaling repressor FaRGA1 instead of activation of GA biosynthesis. However, no significant differences in *FaRGA1* expression were detected between the control and transgenic lines (**Figures 6A,B**), and therefore, the vegetative phenotype of *S*OE lines cannot be explained by neither increased endogenous active GAs levels nor reduced *FaRGA1* expression.

Transcriptional response of flowering time genes to exogenous gibberellin treatment

Our expression studies indicate that the elongated phenotype and increased runner production observed when *FaSOC1* was overexpressed was not due to a rise on GA biosynthesis as previously reported in *F. vesca* (Mouhu et al., 2013). Alternatively, we hypothesized that *FaSOC1* could be a target gene of GA signaling in *F.* × *ananassa* as it has been shown in *Arabidopsis* (Moon et al., 2003). To test this possibility, *F.* × *ananassa* plants were sprayed with 50 ppm GA_3 and gene expression analysis was conducted on GA_3- and mock-treated plants. Since there is strong evidence for negative feedback control of the expression of *GA20ox* and *GA3ox* genes by GA (Hedden and Phillips, 2000), the expression of *FaGA20ox4* and *FaGA3ox1* was analyzed. In our conditions, *FaGA20ox4* expression did not respond to GA application but *FaGA3ox1* was notably down-regulated in GA_3-treated samples, both in leaves and crowns, indicating that *FaGA3ox1* is involved in the negative feedback regulation of the GA biosynthetic pathway in *F.* × *ananassa* (**Supplementary Figure 5**). Despite plants were responding to GA treatment, identical *FaSOC1* transcript levels were found in control and treated plants (**Figure 7A**).

To further investigate the molecular events downstream of GA perception, the expression of other flowering time genes was examined. Transcript accumulation of *FaLFYa*, *FaTFL1*, *FaFT1,* and *FaFT2* did not significantly differ between treated and untreated samples (**Figure 7A**). In contrast, we observed significant differences in the accumulation of *FaCO*, *FaTEM,* and *FaFT3* (**Figure 7A**). In leaves, *FaCO* and *FaTEM* expression followed opposite regulation, with GA_3 mimicking the effect of LD conditions on their expression levels (**Figures 1B,7A**). *FaCO* levels raised after GA_3 treatment whereas *FaTEM* dropped in the same samples. In crowns, a two-fold induction of *FaFT3* was observed in response to GA_3, suggesting that *FaFT3* upregulation in the AXMs might be required for the transition to the vegetative phase in *F.* × *ananassa*.

Gibberellin metabolism genes are differentially regulated in seasonal flowering and perpetual flowering cultivars in response to photoperiod

In strawberry, active GA levels, together with photoperiod and temperature, regulate the transition to the vegetative phase. In *F. vesca*, the signaling cascade mediating these effects has been partially elucidated, with LD upregulating *FvSOC1* which, in turn, induces *FvGA3ox* and *FvGA20ox* gene expression, resulting in higher GA accumulation (Mouhu et al., 2013).

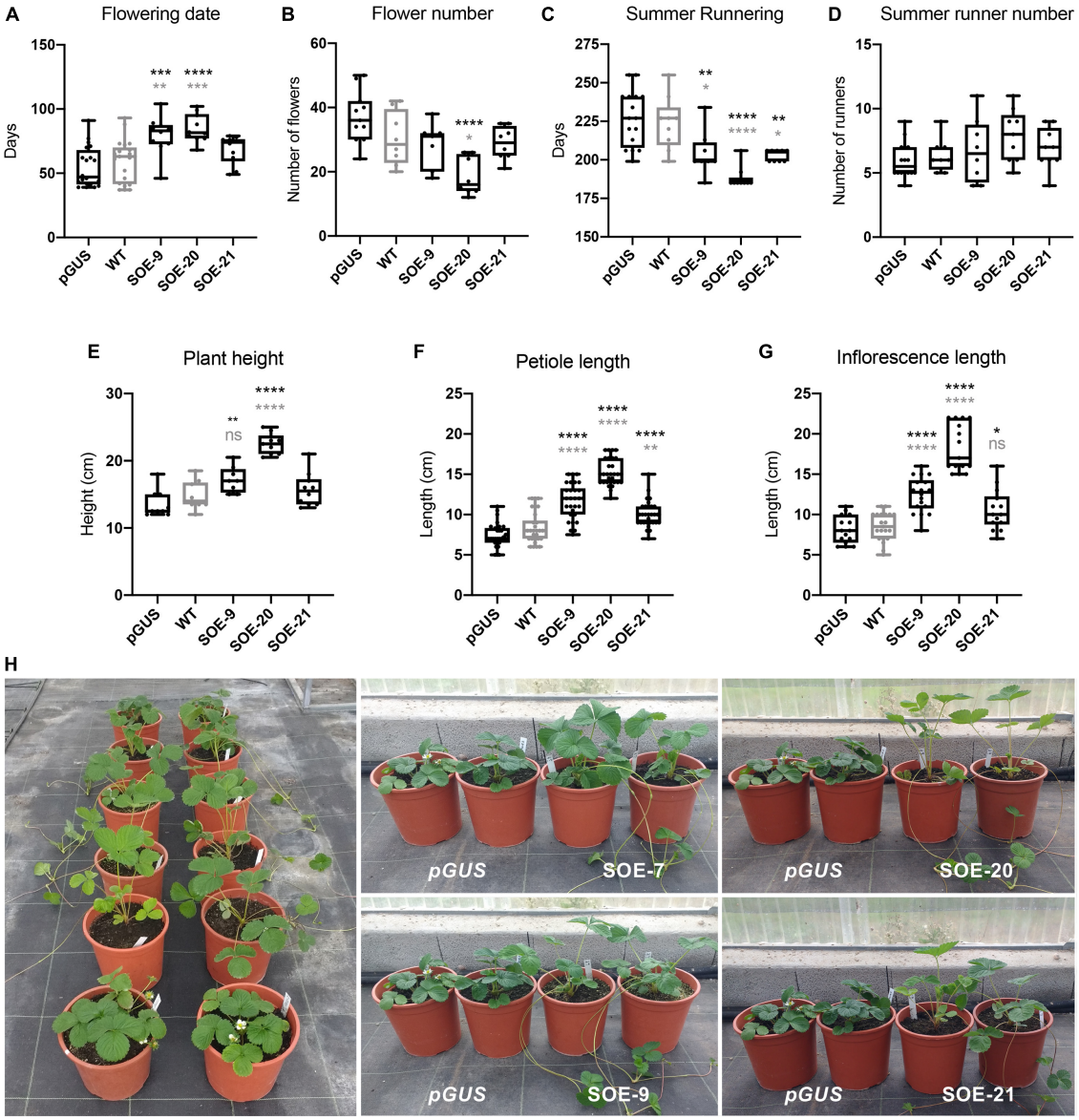


FIGURE 4
Phenotype of 35S:*FaSOC1* lines (SOE) in comparison to control plants. **(A)** Number of days to flowering from November 20th. **(B)** Total flower number (from November 20th until September 1st). **(C)** Number of days until the first summer stolon. **(D)** Number of stolons until September 20th. **(E)** Plant height. **(F)** Length of leaf petioles. **(G)** Length of inflorescence peduncle. Boxes span the 25th and 75th percentiles and the middle line represents the median. Whiskers (T-bars) are the minimum and maximum values. Asterisks represent significant differences between pGUS (in black) or 'Camarosa' wild type (WT; in gray) and 35S:*FaSOC1* over-expresser plants; Statistical significance was determined by ANOVA and Tukey's test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$). **(H)** Phenotype of 35S:*FaSOC1* lines in comparison to control GUS plants. Two representative replicates are shown. Pictures were taken in November.

In *F.* × *ananassa*, GA biosynthesis is required to induce runner production under LD (Hytönen et al., 2009) but, according to our results, its activation is not mediated solely by FaSOC1. Still, GA metabolism genes might be photoperiodically regulated through a FaSOC1 independent pathway. In order to explore how photoperiod regulates GA metabolism in *F.* × *ananassa* and putative differences between cultivars with contrasting flowering habits, *FaGA3ox1, FaGA20ox4,* and *FaGA2ox1* expression was analyzed in 'Selva' and 'Chandler' under SD and LD natural conditions.

FaGA3ox1 expression in leaf and crown was promoted by SD (**Figure 7B**). By contrast, *FaGA20ox4,* which in *F. vesca* is essential for runner production (Tenreira et al., 2017), showed opposite regulation. In leaves, *FaGA20ox4* transcripts were barely detected under any of the photoperiods tested. In SD crowns, *FaGA20ox4* was only detected in one biological replicate

10.3389/fpls.2022.971846

from the SF cultivar Chandler (Figure 7B). However, LD conditions considerably upregulated *FaGA20ox4* accumulation in crowns from the two cultivars. Interestingly, in parallel to *FaGA20ox4* induction in LD crowns, a sharp increase in *FaGA2ox1* expression was detected exclusively in the PF cultivar Selva (Figure 7B). This *FaGA2ox1* upregulation merely in the PF cultivar suggests a putative mechanism to regulate continuous flowering in *F.* × *ananassa*, as *FaGA20ox4* induction in LD-grown meristems would be likely counteracted by the high level of *FaGA2ox1* expression, and therefore GA deactivation would neutralize its biosynthesis.

Discussion

Although the molecular network regulating flowering transition in annual plants, *Arabidopsis* in particular, is well-understood, there has been little research on the molecular mechanisms leading to flower induction in perennial species. Within the past decade, a significant advance in the understanding of the regulation of flowering and runnering in the perennial Rosaceae *F. vesca* has been made, showing substantial differences regarding the role of particular transcription factors compared to what we know from *Arabidopsis* and other annual plants. One of such singularities is the function of the MADS-box transcription factor SOC1. Whereas in annual plants, SOC1 functions as an integrator promoting floral transition and the development of floral organs (Lee and Lee, 2010), strawberry FvSOC1 represses flowering and promotes vegetative growth (Mouhu et al., 2013). The environmental control of flowering in woodland and cultivated strawberry is comparable and therefore, due to the complex genetics of cultivated strawberry, most of the basic knowledge on the molecular events leading to flower induction comes from studies in *F. vesca*, whereas in the octoploid species it is still scarce. The aim of this study was to further advance in the understanding of the genetic control of flowering and runnering in *F.* × *ananassa* and elucidate how well *F. vesca* knowledge is translated to the cultivated strawberry.

FaSOC1 has a conserved role as a repressor of flowering in *F.* × *ananassa*

FaSOC1 expression is detected in a range of vegetative and reproductive tissues, although expression levels are notably lower in flowers and fruits indicative of a major role during vegetative development. No significant diurnal rhythm was observed in its expression neither in SD nor in LD. In contrast, *FaSOC1* expression showed a marked response to day length in both leaf and crown and in 'Chandler' and 'Selva,' with an average six-fold induction under long photoperiods, again supporting its role during the vegetative phase as described

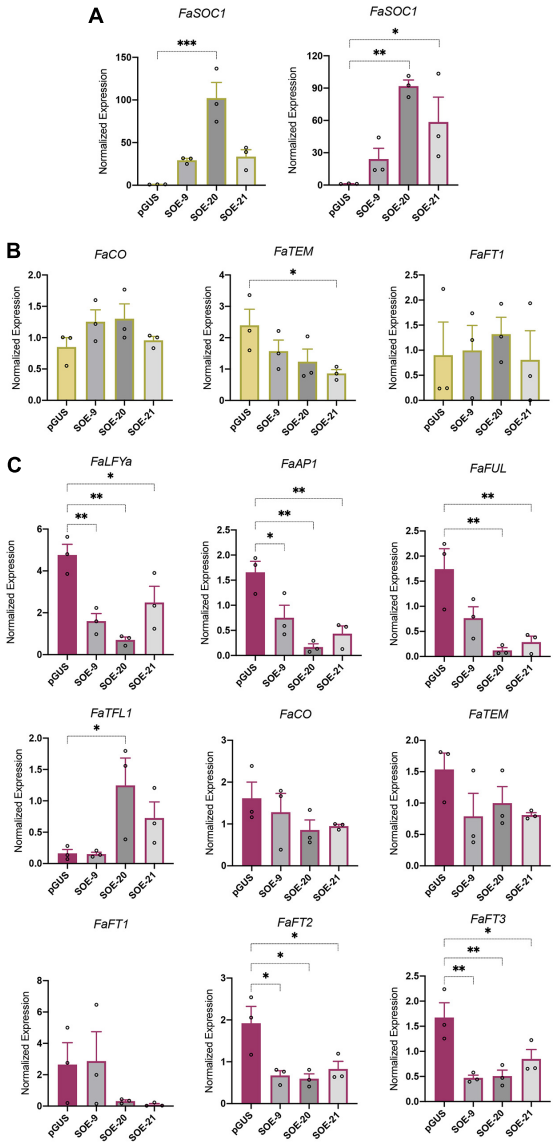


FIGURE 5
Gene expression analysis in 35S:*FaSOC1* lines (SOE). **(A)** *FaSOC1* overexpression in leaves (green, left panel) and crown (purple, right panel) of three independent SOE lines. **(B)** *FaCO, FaTEM and FaFT1* expression in leaves of SOE lines. **(C)** Flowering time gene expression in crowns from SOE lines. All samples were collected under natural SD (9 h52 min/14 h17 min day/night; average maximum and minimum temperature from previous 4 weeks: 21.1/10.6°C) conditions at ZT3. Graphs show the average of normalized values from three biological replicates ± SEM. In each bar, dots represent the normalized value of each biological replicate. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

for *F. vesca FvSOC1* (Mouhu et al., 2013). Overexpression of *FaSOC1* in the SD cultivar Camarosa delayed flowering, confirming the role of FaSOC1 as a strong repressor of flowering. At the transcriptional level, *FaSOC1* overexpression downregulates the expression of *FaFT2* and *FaFT3* in crowns,

with a concomitant decline in the accumulation of floral identity genes such as *FaAP1*, *FaLFYa,* and *FaFUL* (**Figure 5B**). Mouhu et al. (2013) showed FvSOC1 activates the transcription of *FvTFL1* and that FvTFL1 function is required for the FvSOC1-induced flower repression in *F. vesca*. In our SOE lines, *FaTFL1* upregulation was not consistent in all transgenic lines (**Figure 5B**), being only detected in SOE-20 and -21, the ones with higher transgene expression level. A yet uncharacterized SOC1 independent mechanism of *TFL1* regulation has been proposed in *F. vesca* acting at low (<11°C) and high temperatures (>23°C) (Rantanen et al., 2015) or in *F. nilgerrensis* at 11–18°C (Fan et al., 2022), suggesting this other temperature-responsive pathway could be acting in 'Camarosa' in our growing conditions. Since the late flowering phenotype was similar in the three transgenic lines (**Figure 4A**) and it did not correlate with *FaTFL1* expression level, we concluded that in 'Camarosa,' *FaSOC1*-dependent *FaFT2* and *FaFT3* repression in crowns might have a major role mediating flowering repression. Additionally, *FaTFL1* activation is only achieved at higher *FaSOC1* transgene levels and do not correlate with the strength of the flowering time phenotype. These results highlight once more the importance of the TFL1/FT balance in the control of meristem fate in strawberry (Gaston et al., 2021) and suggest FaSOC1 modifies this balance to promote the vegetative state in *F.* × *ananassa* mainly through *FaFT2* and *FaFT3* downregulation.

FaSOC1 promotion of vegetative development occurs independently of gibberellin biosynthesis

In addition to delaying flowering, overexpression of *FaSOC1* has various effects on growth and development during the vegetative phase. In SF octoploid cultivars, under natural growing conditions, a sharp upregulation of *FaSOC1* expression in leaves and crowns correlates with the induction of runner formation and petiole elongation, which are normal photoperiodic responses to increasing day lengths (Guttridge and Thompson, 1964). We show here that *FaSOC1* is involved in the regulation of both processes as petioles are significantly longer and the emergence of runners takes place 3–6 weeks earlier in 35S:*FaSOC1* than in control lines, although the final count of runners is only slightly increased at the end of the season (**Figure 4**). Similar morphological and physiological effects are observed after exogenous GA treatment (Thompson and Guttridge, 1959; Guttridge and Thompson, 1964; Hytönen et al., 2009). In fact, the phenotype of *F. vesca* plants overexpressing *FvSOC1*, comparable to that in our 35S:*FaSOC1* lines, is due to an upregulation of *GA3-ox* and *GA20-ox* genes, leading to increased GA accumulation (Mouhu et al., 2013). However, we could not confirm a similar mechanism in 'Camarosa' SOE lines, as none of the tested genes in the GA metabolic pathway presented differential expression compared to the control (**Figure 6**). The activity of the GA biosynthetic enzyme GA20ox is of particular importance in determining GA concentration in many plant species. In *F. vesca, FvGA20ox4* activation in AXM under LD has been shown to be required for stolon development. Whereas at 18°C this activation occurs *via* an FvSOC1-dependent photoperiodic pathway (Mouhu et al., 2013; Tenreira et al., 2017; Andrés et al., 2021), *FvGA20ox4* upregulation also occurs independently of FvSOC1 (Andrés et al., 2021; Fan et al., 2022). The latter situation is therefore in line with our observations in cultivated strawberry. Although *FaSOC1* overexpression did not affect that of *FaGA20ox4*, this gene showed a marked photoperiodic response in crowns in response to LD (**Figure 7B**), suggesting it plays a critical role in the accumulation of active GA required for inducing the vegetative phase, in agreement with its role in the diploid species. Furthermore, an interesting difference between the PF cultivar Selva and the SF cultivar Chandler was observed in *FaGA2ox1* expression, coding for a GA inactivation enzyme (**Figure 7B**). *FaGA2ox1* transcript accumulation in 'Chandler' was low under SD and LD photoperiods, and equivalent to that in 'Selva' under SD. However, a marked upregulation of *FaGA2ox1* was detected in 'Selva' under LD. Increased FaGA2ox1 activity is expected to decrease the levels of bioactive GAs, negatively affecting runner production and promoting branch crown development, aiding continuous flowering in 'Selva.'

In an attempt to explain the constitutive GA response in SOE lines, we considered the possibility of *FaSOC1* being a mediator of GA signaling, activated downstream of GA perception, similar to Arabidopsis *SOC1*, which is induced by GAs (Lee and Lee, 2010). We could not confirm this hypothesis as, in our studies, *FaSOC1* expression was equivalent in GA-treated and mock plants (**Figure 7A**). However, although plants were sensing and responding to GA, a weakened response to exogenous GA cannot be completely ruled out under our photoperiodic conditions, as daylight was slightly below 12h at the time of the first application. Previous studies showed *F.* × *ananassa* plants are able to respond to exogenous GA treatment under SD (Guttridge and Thompson, 1959), but also a reduced sensibility has been observed in short compared to long photoperiods (Hytönen et al., 2009).

Considering *FaSOC1* expression does not respond to GA treatment, we therefore hypothesized that GA signaling and FaSOC1 might have common target genes that would account for the vegetative phenotype observed in SOE lines. Among the flowering genes evaluated in this work, *FaTEM1* expression is downregulated in leaves of GA treated and 35S:*FaSOC1* plants. Notably, a similar trend was observed in 'Selva' and 'Chandler' grown under LD (**Figure 1B**). The fact that *FaTEM* was downregulated in different scenarios, all of them promoting vegetative development, pointed to a putative role of this factor in the repression of flowering in *F.* × *ananassa*. TEM proteins

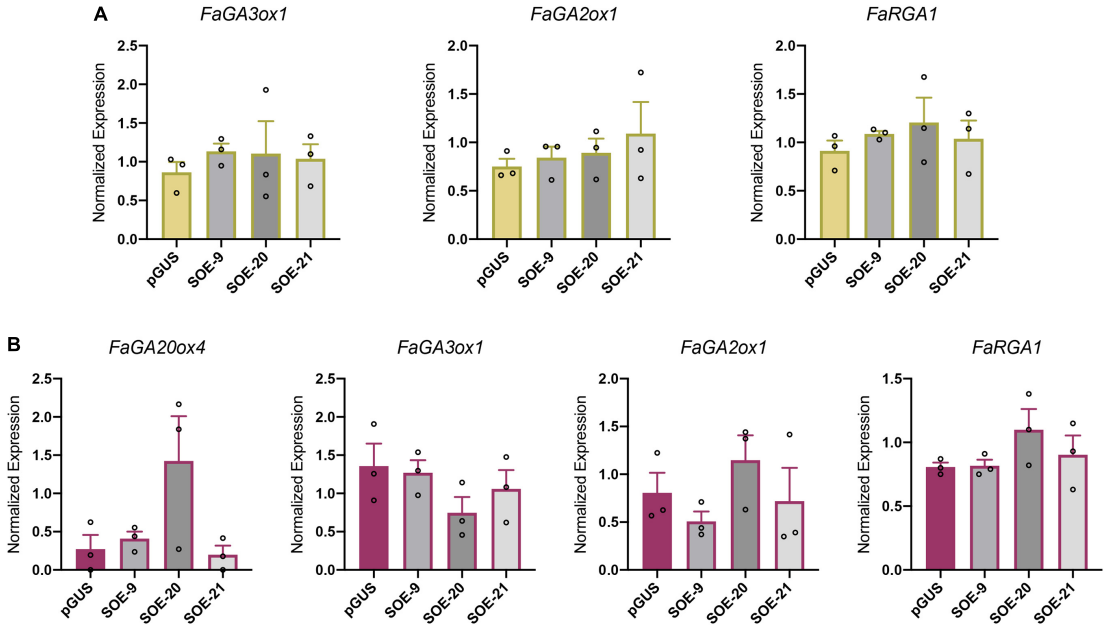


FIGURE 6
Overexpression of *FaSOC1* does not alter the expression of GA metabolism and signaling related genes. **(A)** *FaGA3ox1*, *FaGA2ox1*, and *FaRGA1* expression in leaves of three independent SOE lines. **(B)** *FaGA20ox4*, *FaGA3ox1*, *FaGA2ox1*, and *FaRGA1* expression in crowns from SOE lines. All samples were collected under SD. Graphs show the average of normalized values from three biological replicates ± SEM. In each bar, dots represent the normalized value of each biological replicate.

have been proposed to act as regulators of the juvenile vegetative phase in species as diverse as antirrhinum, *Arabidopsis* or olive (Sgamma et al., 2014), including another perennial Rosaceae such as loquat (Peng et al., 2021). In *Arabidopsis, TEM1* represses photoperiodic flowering at early stages of vegetative growth, then its expression levels progressively decay throughout development to reach a minimum at the time of floral transition, thus allowing *FT* activation by CO (Castillejo and Pelaz, 2008). Recently, it has been shown that heterologous overexpression of apple *MdTEM* in *F. vesca* H4 leads to *FvFT1* induction, whereas the opposite effect is observed in RNAi-*MdTEM1* lines (Dejahang et al., 2022). In *35S:MdTEM1* and *35S:MdTEM2* lines, generated in the *fvtfl1* mutant background of H4, elevated *FvFT1* expression levels in leaves correlated with an extremely early flowering (Dejahang et al., 2022), phenocopying *FvFT1* overexpression in H4 (Rantanen et al., 2014). However, in a *TFL1* WT genotype, *FT1* upregulation in leaves would be expected to induce TFL1 in the SAM and repress flowering.

FaCO overexpression does not prevent flowering under short days in *F.* × *ananassa*

In *F. vesca* H4, FvCO is required for *FvFT1* induction in leaves in response to long photoperiods, functioning as a strong promoter of flowering under LD as there is no functional FvTFL1 (Kurokura et al., 2017). However, in SF genotypes carrying a functional TFL1, the role of CO in the photoperiodic control of the alternance between vegetative and reproductive development needs further investigation. According to the model proposed in *F. vesca*, and considering 'Camarosa' harbors a functional allele of the flowering repressor FaTFL1, an expected outcome of *FaCO* overexpression in 'Camarosa' was the activation of the FaFT1-FaSOC1-FaTFL1 module and consequently, flowering repression and promotion of vegetative development. As shown in **Figure 1**, in *F.* × *ananassa*, we observed a correlation in the upregulation of FaCO and FaFT1 in leaf, and FaSOC1 and FaTFL1 in crowns by LDs. However, overexpression of *FaCO* in SF 'Camarosa' only causes a subtle delay on flowering time and a similar discrete advance of runnering (**Figure 3A**), resulting in a slightly extended vegetative phase. These results might indicate the existence of an additional factor that prevents vegetative development under SD even in the presence of elevated FaCO. A putative candidate could be FaTEM, as it has been shown to repress *FaFT1* expression (Dejahang et al., 2022) and it is upregulated in response to SD (**Figure 1A**). As a result, elevated FaTEM expression levels would not allow *FaFT1* activation by FaCO. Additionally, in contrast to the moderate effect of FaCO in prolonging the vegetative phase, a negative effect of *FaCO* overexpression on runner capacity was detected in most of the T_0 transgenic lines and further observed in COE-12 (**Figure 3D**)

in the following season. Although T_0 phenotypes might be affected by the fact that those plants come directly from *in vitro* culture and maybe not related to transgene expression, the possibility of FaCO negatively affecting runner initiation should not be discarded and, in that case, the phenotypes of the three selected runnering lines might not fully reflect FaCO function. In depth characterization of 35S:*FaCO* transgenic lines awaits in order to elucidate the molecular mechanisms regulating FaCO effects on flowering and runnering.

Expression analysis of *FaFT* genes suggest functional diversification and specific roles in the regulation of flowering and vegetative development

In this work we propose a role for FaCO and FaSOC1 as weak and strong repressors, respectively, of the photoperiodic flowering pathway in *F.* × *ananassa*. Additionally, our expression analysis performed in the SD and PF cultivars, 'Chandler' and 'Selva' respectively, shed some light into the molecular mechanisms leading to flower induction in both of them. In many different species, *FT* genes have been shown to be major components of the florigen, a graft-transmissible signal produced in the leaves that induces flowering at the shoot apex in response to inductive photoperiods (Andrés and Coupland, 2012). In addition to regulating flowering, FT genes have been implicated in a range of physiological processes, such as promotion of vegetative development in poplar (Böhlenius et al., 2006), tuberization or bulbing in potato and onion, respectively (Navarro et al., 2011; Lee et al., 2013), or more pleiotropic roles as general regulators of growth, like in tomato or maize (Lifschitz et al., 2006; Danilevskaya et al., 2011). In the strawberry genome, three *FT* homologs have been identified (Nakano et al., 2015), however there is still scarce knowledge about the physiological process controlled by each FaFT member, or if variations in their expression pattern are able to explain differences in the vegetative and flowering responses of SD or PF cultivars. Most recent studies overexpressing *F. vesca FvFT2* in *F. vesca* and *F.* × *ananassa* suggest that FaFT2 might act as florigen under SD in cultivated strawberry (Gaston et al., 2021).

In SF cultivars like 'Chandler,' floral induction usually takes place at late fall, when diurnal light periods and temperatures become lower, whereas in PF cultivars as 'Selva,' flowers are initiated continuously throughout the growing season from spring until late autumn. The length of the flowering period is also variable among cultivars within the SF or PF habits and dependent on genetic and environmental conditions (Stewart and Folta, 2010; Heide et al., 2013; Labadie et al., 2019). In the present study, we have conducted an expression analysis of the three *FaFT* homologs in crowns and leaves of *F.* × *ananassa* 'Selva' and 'Chandler' grown under natural conditions and collected in SD (December 21st) or LD (June 21st). Under floral promoting SD, higher expression levels of the three *FaFT* homologs, with the exception of *FaFT2* in 'Selva,' are observed in crowns, and correlate with the induction of floral identity genes such as *FaAP1* and *FaLFY* and downregulation of the floral repressor *FaTFL1* (**Figures 1A,B**). *FaFT1* and *FaFT3* upregulation in shoot tips under floral inductive conditions have been previously described in other octoploid cultivars (Nakano et al., 2015; Koembuoy et al., 2020). SD specific *FaFT2* induction in shoot tips was also documented by Nakano et al. (2015) after 7 days of photoperiod treatments, although no differences were observed after 21 days. Interestingly, *FaFT2* and *FaFT3* induction in crowns is abolished when *FaSOC1* is overexpressed (**Figure 5B**), suggesting SD dependent *FaSOC1* downregulation (**Figure 1**) is required to allow activation of the floral promoter *FaFT* genes. Alternatively, FaSOC1 might directly repress FaFT2 and FaFT3 in meristems.

Gene expression changes in crowns are expected to occur in response to a long-distance signal coming from leaves (Colasanti and Sundaresan, 2000; Andrés and Coupland, 2012). Accordingly, we observed SD photoperiods induce the accumulation of *FaFT2* and *FaFT3* transcripts in leaves, being more evident in SF 'Chandler' than in PF 'Selva' at the time sampling was performed (**Figure 1B**). Upregulation of *FvFT3/FaFT3* by SD was not detected in leaves in previous studies by Nakano et al. (2015) or Gaston et al. (2021). In this latter study, *FaFT2* expression was equivalent in SD and LD leaf samples from the *F.* × *ananassa* Japanese cultivar Nyoho. In contrast, Gaston et al. (2021) detected a transient upregulation of *FvFT2* under SD conditions in a PF, but not in a SF, *F. vesca* genotype. In their monthly time-course performed from June to November under natural environmental conditions, *FvFT2* peaked in October. The observed discrepancies might be due to natural variation among genotypes. Alternatively, peaks might be easily missed due to their transient nature. *FaFT2* induction in PF 'Selva' could have already happened when we sampled in December, whereas *FvFT2* rise in the SF cultivar used in Gaston et al. (2021) could have taken place after their last sample in November. According to our qPCR data, upregulation of *FaFT2* and *FaFT3* in leaf is compatible with both of them having partially overlapping roles as florigens. Indeed, *FvFT2* has been already shown to function as a strong floral promoter when overexpressed in *F. vesca* and *F.* × *ananassa* (Gaston et al., 2021; Sabbadini et al., 2021). In the same study, *FvFT3* overexpression only caused a modest advance of flowering in *F. vesca*. In contrast, when *FvFT3* is overexpressed in *F.* × *ananassa,* it promotes the vegetative state of AXM and thus runner development (Gaston et al., 2021; Sabbadini et al., 2021). This apparently contradictory result, with *FaFT3* expression in leaves associated with floral inductive conditions, and the promotion of runner production when overexpressed, raised the possibility of FaFT3 fulfilling different roles, as floral inducer in leaves and SAM and as runner promoter in AXM. In *F.* × *ananassa*, GA is a

10.3389/fpls.2022.971846

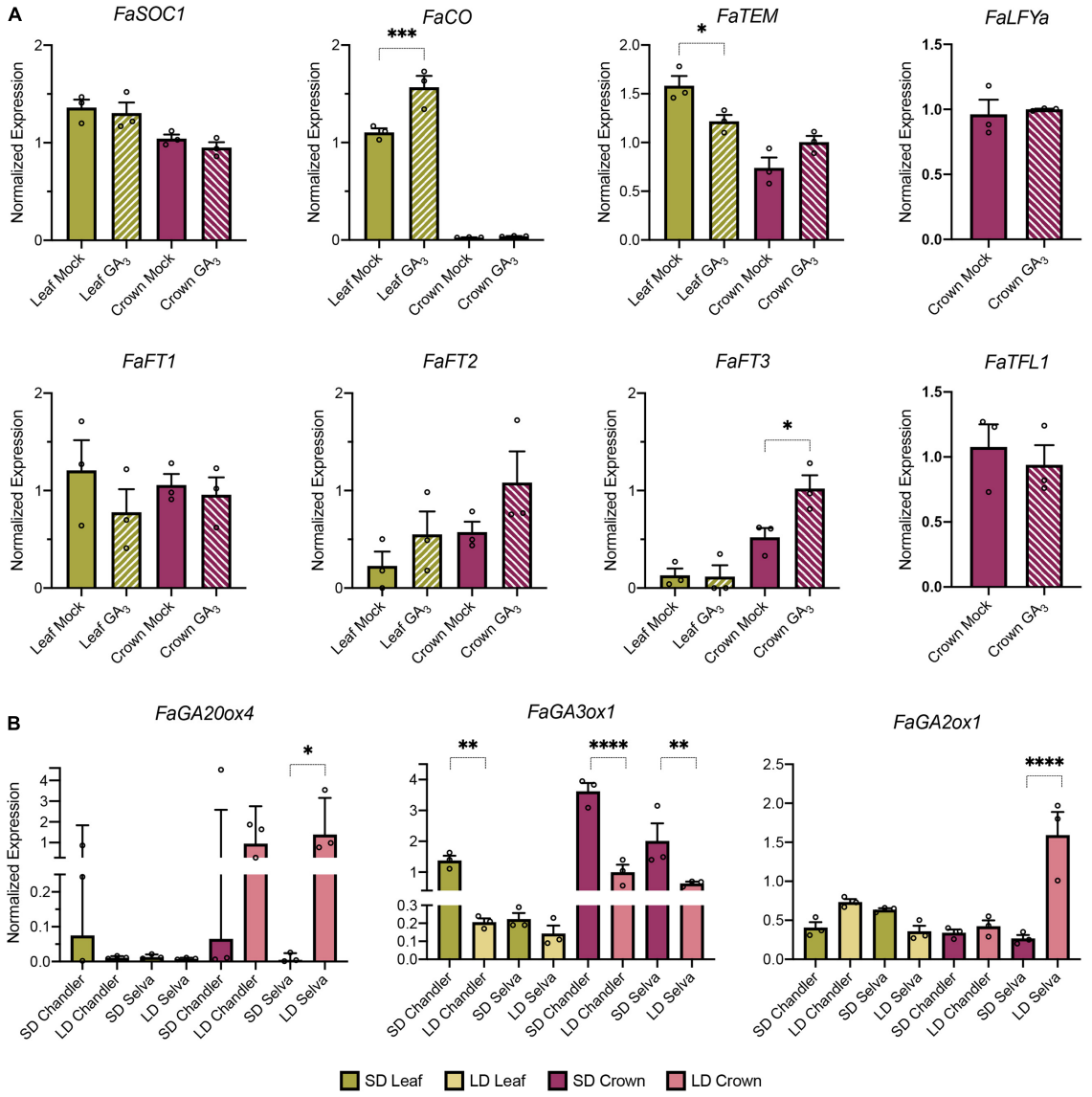


FIGURE 7
Role of GA in the shift between floral and vegetative development in *F.* × *ananassa*. **(A)** Flowering time gene expression in response to exogenous GA_3 treatment in leaves (green) and crowns (purple) from SF 'Camarosa' plants. **(B)** GA metabolism gene expression in leaves and crowns from the SF cultivar Chandler and the PF cultivar Selva collected under SD and LD. Graphs show the average of normalized values from three biological replicates ± SEM. In each bar, dots represent the normalized value of each biological replicate. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

key signal determining bud fate and needed for runner initiation (Hytönen et al., 2009). Therefore, we quantified *FaFT1-FaFT3* expression in GA-treated samples. Interestingly, *FaFT1* and *FaFT2* expression do not respond to GA, but a two-fold *FaFT3* induction is observed specifically in crowns after exogenous GA application (**Figure 7A**), supporting a key role for FaFT3 in the regulation of the vegetative developmental fate of AXM downstream GA perception.

As for *FaFT1*, its expression pattern follows opposite photoperiodic regulation in crowns compared to leaves, suggesting it might also fulfill tissue specific roles. Similarly to *F. vesca FvFT1* (Koskela et al., 2012), *FaFT1* is induced in leaves under LD, promoting the vegetative state of plant meristems. However, under SD, when *FaTFL1* expression drops in the meristems, *FaFT1* can act as floral inducer in crowns as previously shown in *F. vesca* H4 in LD or when *FvFT1* is overexpressed (Rantanen et al., 2014).

Altogether, our results suggest the three *F.* × *ananassa* FT proteins have overlapping and particular roles in the regulation of the alternance between flowering and vegetative

development. FaFT2 is proposed as the florigenic signal transiently produced in leaves under SD, which likely exerts its function redundantly with FaFT3. At the SAM, *FaFT1-3* induction under SD indicates overlapping roles in the promotion of flowering transition. Additionally, the transition to the vegetative phase would require upregulation of FaFT1 and FaFT3 in leaf and AXMs, respectively.

Overall, this work supports that *F.* × *ananassa* FaSOC1 has retained FvSOC1 role as repressor of flowering and promoter of vegetative development. However, signaling events downstream FaSOC1 show particularities compared to *F. vesca* that hinders direct transfer of knowledge acquired in the diploid species to cultivated ones. On the contrary, FvCO has been shown to promote flowering while repressing runner development in *fvtfl1 F. vesca* accessions, whereas our study suggests that *FaCO* responds to LD photoperiods as in *F. vesca* with a different outcome, as flowering was marginally delayed in COE lines. Comparison of FvCO and FaCO protein sequences did not reveal substantial differences able to explain the contrasting phenotypes in the two *Fragaria* species. A total of 13 aminoacid substitutions were found in the coding sequence, 6 non-synonymous, but none of them affected conserved residues (**Supplemental Figure 6**). The presence of a functional *FaTFL1* allele in the SF cultivar Camarosa and COE lines is an obvious difference with reported studies in F. vesca and if its activation occurs downstream of FaCO, as it has been suggested for FvTFL1, a more dramatic repression of flowering was expected. The subtle negative effect on flowering of *FaCO* overexpression and the contrasting effects on runner initiation suggest the existence of putative additional factors that counterbalance its activity, particularly under SD conditions.

Moreover, our results provide a groundwork for detailed characterization of factors that may play a critical role in the tradeoff between flowering and runnering in *F.* × *ananassa*. One of such factors is FaTEM, which expression pattern is compatible with a role as repressor of vegetative development under SD, and thus promoter of flowering, likely through preventing FaCO-dependent *FaFT1* activation in leaves. Additionally, LD induction of *FaGA2ox1* in crowns specifically in PF 'Selva' suggests a putative mechanism to lower active GA levels in PF genotypes in order to enable successive rounds of flowering. Lastly, detailed expression of *FaFT* genes suggest the three homologs from *F.* × *ananassa* might have overlapping roles inducing flowering in the apical meristem under SD, but also functional divergence in the case of FaFT1 and FaFT3 is proposed. Although further studies are needed to clarify these mechanisms, our work broadens our knowledge on the regulatory pathways controlling flowering and vegetative growth in cultivated strawberry in response to photoperiod, and how these compare with flowering regulation in annual plants or the diploid relative *F. vesca*.

Data availability statement

The data presented in this study are deposited in the GenBank repository, accession numbers ON996971 *(FaCO)* and ON996972 *(FaSOC1)*.

Author contributions

JM-A, CC, and IA designed the experiments. JS-S provided plant materials and participated in the design of experimental work. JM-A, CP, and CC carried out the experimental work and statistical analyses. IA and CC supervised the work and wrote the manuscript. All authors commented on and accepted the final manuscript version.

Funding

This study was supported by IFAPA and FEDER (PR.AVA.AVA2019.034), the Spanish Research Agency (PID111496RR-I00/AEI/10.13039/501100011033), and Junta de Andalucía and FEDER (P18-RT-4856).

Acknowledgments

We are grateful to Francisco J. Durán for his excellent care of strawberry plants.

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.971846/full#supplementary-material

References

Andrés, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nat. Rev. Genet.* 13, 627–639. doi: 10.1038/nrg3291

Andrés, J., Caruana, J., Liang, J., Samad, S., Monfort, A., Liu, Z., et al. (2021). Woodland strawberry axillary bud fate is dictated by a crosstalk of environmental and endogenous factors. *Plant Physiol.* 187, 1221–1234. doi: 10.1093/plphys/kiab421

Black, B. L. (2004). Prohexadione-calcium decreases fall runners and advances branch crowns of 'Chandler' strawberry in a cold-climate annual production system. *J. Am. Soc. Horticult. Sci.* 129, 479–485. doi: 10.21273/JASHS.129.4.0479

Böhlenius, H., Huang, T., Charbonnel-Campaa, L., Brunner, A. M., Jansson, S., Strauss, S. H., et al. (2006). CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science* 312, 1040–1043. doi: 10.1126/science.1126038

Bradford, E., Hancock, J. F., and Warner, R. M. (2010). Interactions of temperature and photoperiod determine expression of repeat flowering in strawberry. *J. Amer. Soc. Hort. Sci.* 135, 102–107. doi: 10.21273/JASHS.135.2.102

Brown, T., and Wareing, P. F. (1965). The genetical control of the everbearing habit and three other characters in varieties of *Fragaria vesca*. *Euphytica* 14, 97–112. doi: 10.1007/BF00032819

Caruana, J. C., Sittmann, J. W., Wang, W., and Liu, Z. (2018). Suppressor of runnerless encodes a DELLA protein that controls runner formation for asexual reproduction in strawberry. *Mol. Plant* 11, 230–233. doi: 10.1016/j.molp.2017.11.001

Castillejo, C., and Pelaz, S. (2008). The balance between CONSTANS and TEMPRANILLO activities determines FT expression to trigger flowering. *Curr. Biol.* 18, 1338–1343. doi: 10.1016/j.cub.2008.07.075

Castro, P., Bushakra, J. M., Stewart, P. J., Weebadde, C. K., Wang, D., Hancock, J. F., et al. (2015). Genetic mapping of day-neutrality in cultivated strawberry. *Mol. Breed.* 35, 79–95. doi: 10.1007/s11032-015-0250-4

Colasanti, J., and Sundaresan, V. (2000). 'Florigen' enters the molecular age: long-distance signals that cause plants to flower. *Trends Biochem. Sci.* 25, 236–240. doi: 10.1016/s0968-0004(00)01542-5

Danilevskaya, O. N., Meng, X., McGonigle, B., and Muszynski, M. G. (2011). Beyond flowering time: pleiotropic function of the maize flowering hormone florigen. *Plant Signal Behav.* 6, 1267–1270. doi: 10.4161/psb.6.9.16423

Darrow, G. M. (1966). *The Strawberry. History, Breeding and Physiology*. New York, NY: Holt, Rinehart and Winston.

Dejahang, A., Maghsoudi, N., Mousavi, A., Farsad-Akhtar, N., Hernandez, L. M., Pelaz, S., et al. (2022). TEMPRANILLO homologs in apple regulate flowering time in the woodland strawberry *Fragaria vesca*. *bioRxiv [preprint]* doi: 10.1101/2022.02.08.479567

Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., et al. (2019). Origin and evolution of the octoploid strawberry genome. *Nat. Genet.* 51, 541–547. doi: 10.1038/s41588-019-0356-4

El Mansouri, I., Mercado, J. A., Valpuesta, V., López-Aranda, J. M., Pliego-Alfaro, F., and Quesada, M. A. (1996). Shoot regeneration and Agrobacterium-mediated transformation of *Fragaria vesca* L. *Plant Cell Rep.* 15, 642–646. doi: 10.1007/BF00232469

Fan, G., Andrés, J., Olbricht, K., Koskela, E., and Hytönen, T. (2022). Natural variation in the control of flowering and shoot architecture in diploid fragaria species. *Front. Plant Sci.* 13:832795. doi: 10.3389/fpls.2022.832795

Gambino, G., Perrone, I., and Gribaudo, I. (2008). A Rapid and effective method for RNA extraction from different tissues of grapevine and other woody plants. *Phytochem. Anal.* 19, 520–525. doi: 10.1002/pca.1078

Gaston, A., Perrotte, J., Lerceteau-Kohler, E., Rousseau-Gueutin, M., Petit, A., Hernould, M., et al. (2013). PFRU, a single dominant locus regulates the balance between sexual and asexual plant reproduction in cultivated strawberry. *J. Exp. Bot.* 64, 1837–1848. doi: 10.1093/jxb/ert047

Gaston, A., Potier, A., Alonso, M., Sabbadini, S., Delmas, F., Tenreira, T., et al. (2021). The FveFT2 florigen/FveTFL1 antiflorigen balance is critical for the control of seasonal flowering in strawberry while FveFT3 modulates axillary meristem fate and yield. *New Phytol.* 232, 372–387. doi: 10.1111/nph.17557

Guttridge, C., and Thompson, P. (1964). The effect of gibberellins on growth and flowering of Fragaria and Duchesnea. *J. Exp. Bot.* 15, 631–646. doi: 10.1093/jxb/15.3.631

Guttridge, C. G. (1958). The Effects of winter chilling on the subsequent growth and development of the cultivated strawberry plant. *J. Horticult. Sci.* 33, 119–127. doi: 10.1080/00221589.1958.11513920

Guttridge, C. G. (1985). *CRC Handbook of Flowering Vol. III Fragaria x Ananassa*. New York, NY: CRC Press.

Guttridge, C. G., and Thompson, P. A. (1959). Effect of gibberellic acid on length and number of epidermal cells in petioles of strawberry. *Nature* 183, 197–198. doi: 10.1038/183197b0

Hayama, R., Yokoi, S., Tamaki, S., Yano, M., and Shimamoto, K. (2003). Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* 422, 719–722. doi: 10.1038/nature01549

Hedden, P., and Phillips, A. L. (2000). Gibberellin metabolism: new insights revealed by the genes. *Trends Plant Sci.* 5, 523–530. doi: 10.1016/s1360-1385(00)01790-8

Heide, O. M. (1977). Photoperiod and temperature interactions in growth and flowering of strawberry. *Physiol. Plant.* 40, 21–26. doi: 10.1111/j.1399-3054.1977.tb01486.x

Heide, O. M., Stavang, J. A., and Sønsteby, A. (2013). Physiology and genetics of flowering in cultivated and wild strawberries – a review. *J. Horticult. Sci. Biotechnol.* 88, 1–18. doi: 10.1080/14620316.2013.11512930

Hossain, M. R., Natarajan, S., Kim, H.-T., Jesse, D. M. I., Lee, C.-G., Park, J.-I., et al. (2019). High density linkage map construction and QTL mapping for runner production in allo-octoploid strawberry Fragaria × ananassa based on ddRAD-seq derived SNPs. *Sci. Rep.* 9:3275. doi: 10.1038/s41598-019-39808-9

Hytönen, T., Elomaa, P., Moritz, T., and Junttila, O. (2009). Gibberellin mediates daylength-controlled differentiation of vegetative meristems in strawberry (Fragaria × ananassa Duch). *BMC Plant Biol.* 9:18. doi: 10.1186/1471-2229-9-18

Hytönen, T., and Kurokura, T. (2020). Control of flowering and runnering in strawberry. *Horticult. J.* 89, 96–107. doi: 10.2503/hortj.UTD-R011

Hytönen, T., Mouhu, K., Koivu, I., and Junttila, O. (2008). Prohexadione-calcium enhances the cropping potential and yield of strawberry. *Eur. J. Horticult. Sci.* 73, 210–215.

Iwata, H., Gaston, A., Remay, A., Thouroude, T., Jeauffre, J., Kawamura, K., et al. (2012). The TFL1 homologue KSN is a regulator of continuous flowering in rose and strawberry. *Plant J.* 69, 116–125. doi: 10.1111/j.1365-313X.2011.04776.x

Jonkers, H. (1965). *On the Flower Formation, the Dormancy and the Early Forcing of Strawberries*. Amsterdam: Veenman.

Koembuoy, K., Hasegawa, S., Otagaki, S., Takahashi, H., Nagano, S., Isobe, S., et al. (2020). RNA-seq Analysis of Meristem Cells Identifies the FaFT3 Gene as a Common Floral Inducer in Japanese cultivated strawberry. *Horticult. J.* 89, 138–146. doi: 10.2503/hortj.UTD-126

Koskela, E. A., Kurokura, T., Toivainen, T., Sønsteby, A., Heide, O. M., Sargent, D. J., et al. (2017). Altered regulation of TERMINAL FLOWER 1 causes the unique vernalisation response in an arctic woodland strawberry accession. *New Phytol* 216, 841–853. doi: 10.1111/nph.14734

Koskela, E. A., Mouhu, K., Albani, M. C., Kurokura, T., Rantanen, M., Sargent, D. J., et al. (2012). Mutation in TERMINAL FLOWER1 reverses the photoperiodic requirement for flowering in the wild strawberry *Fragaria vesca*. *Plant Physiol.* 159, 1043–1054. doi: 10.1104/pp.112.196659

Koskela, E. A., Sønsteby, A., Flachowsky, H., Heide, O. M., Hanke, M. V., Elomaa, P., et al. (2016). TERMINAL FLOWER1 is a breeding target for a novel everbearing trait and tailored flowering responses in cultivated strawberry (*Fragaria* × *ananassa* Duch.). *Plant Biotechnol. J.* 14, 1852–1861. doi: 10.1111/pbi.12545

Kurokura, T., Samad, S., Koskela, E., Mouhu, K., and Hytönen, T. (2017). Fragaria vesca CONSTANS controls photoperiodic flowering and vegetative development. *J. Exp. Bot.* 68, 4839–4850. doi: 10.1093/jxb/erx301

Labadie, M., Denoyes, B., and Guédon, Y. (2019). Identifying phenological phases in strawberry using multiple change-point models. *J. Exp. Bot.* 70, 5687–5701. doi: 10.1093/jxb/erz331

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* 61, 2247–2254. doi: 10.1093/jxb/erq098

Lee, R., Baldwin, S., Kenel, F., McCallum, J., and Macknight, R. (2013). FLOWERING LOCUS T genes control onion bulb formation and flowering. *Nat. Commun.* 4:2884. doi: 10.1038/ncomms3884

Li, W., Zhang, J., Sun, H., Wang, S., Chen, K., Liu, Y., et al. (2018). FveRGA1, encoding a DELLA protein, negatively regulates runner production in *Fragaria vesca*. *Planta* 247, 941–951. doi: 10.1007/s00425-017-2839-9

Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., et al. (2006). The tomato FT ortholog triggers systemic signals that regulate growth

and flowering and substitute for diverse environmental stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6398–6403. doi: 10.1073/pnas.0601620103

Moon, J., Suh, S. S., Lee, H., Choi, K. R., Hong, C. B., Paek, N. C., et al. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *Plant J.* 35, 613–623. doi: 10.1046/j.1365-313x.2003.01833.x

Mouhu, K., Kurokura, T., Koskela, E. A., Albert, V. A., Elomaa, P., and Hytönen, T. (2013). The Fragaria vesca Homolog of SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 represses flowering and promotes vegetative growth. *Plant Cell* 25, 3296–3310. doi: 10.1105/tpc.113.115055

Nakano, Y., Higuchi, Y., Yoshida, Y., and Hisamatsu, T. (2015). Environmental responses of the FT/TFL1 gene family and their involvement in flower induction in Fragaria × ananassa. *J. Plant Physiol.* 177, 60–66. doi: 10.1016/j.jplph.2015.01.007

Navarro, C., Abelenda, J. A., Cruz-Oró, E., Cuéllar, C. A., Tamaki, S., Silva, J., et al. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* 478, 119–122. doi: 10.1038/nature10431

Nishizawa, T. (1993). The effect of paclobutrazol on growth and yield during first year greenhouse strawberry production. *Sci. Horticult.* 54, 267–274. doi: 10.1016/0304-4238(93)90105-Y

Peng, Z., Wang, M., Zhang, L., Jiang, Y., Zhao, C., Shahid, M. Q., et al. (2021). EjRAV1/2 delay flowering through transcriptional repression of EjFTs and EjSOC1s in Loquat. *Front. Plant Sci.* 12:816086. doi: 10.3389/fpls.2021.816086

Perrotte, J., Gaston, A., Potier, A., Petit, A., Rothan, C., and Denoyes, B. (2016a). Narrowing down the single homoeologous FaPFRU locus controlling flowering in cultivated octoploid strawberry using a selective mapping strategy. *Plant Biotechnol. J.* 14, 2176–2189. doi: 10.1111/pbi.12574

Perrotte, J., Guédon, Y., Gaston, A., and Denoyes, B. (2016b). Identification of successive flowering phases highlights a new genetic control of the flowering pattern in strawberry. *J. Exp. Bot.* 67, 5643–5655. doi: 10.1093/jxb/erw326

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, 2002–2007. doi: 10.1093/nar/29.9.e45

Prisca, M., Maarten, V., Jan, V. D., Bart, N., Wouter, S., Timo, H., et al. (2022). Blue and far-red light control flowering time of woodland strawberry (Fragaria vesca) distinctively via CONSTANS (CO) and FLOWERING LOCUS T1 (FT1) in the background of sunlight mimicking radiation. *Environ. Exp. Bot.* 198:104866. doi: 10.1016/j.envexpbot.2022.104866

Rantanen, M., Kurokura, T., Jiang, P., Mouhu, K., and Hytönen, T. (2015). Strawberry homologue of TERMINAL FLOWER1 integrates photoperiod and temperature signals to inhibit flowering. *Plant J.* 82, 163–173. doi: 10.1111/tpj.12809

Rantanen, M., Kurokura, T., Mouhu, K., Pinho, P., Tetri, E., Halonen, L., et al. (2014). Light quality regulates flowering in FvFT1/FvTFL1 dependent manner in the woodland strawberry *Fragaria vesca*. *Front. Plant Sci.* 5:271.doi: 10.3389/fpls.2014.00271

Sabbadini, S., Gaston, A., Potier, A., Denoyes, B., Cappelletti, R., Giovanetti, G., et al. (2021). Isolation and phenotypical characterization of the FT-like genes in strawberry (*Fragaria* × *ananassa*). *Acta Hortic.* 1309, 217–222. doi: 10.17660/ActaHortic.2021.1309.32

Serrano-Bueno, G., Romero-Campero, F. J., Lucas-Reina, E., Romero, J. M., and Valverde, F. (2017). Evolution of photoperiod sensing in plants and algae. *Curr. Opin. Plant Biol.* 37, 10–17. doi: 10.1016/j.pbi.2017.03.007

Sgamma, T., Jackson, A., Muleo, R., Thomas, B., and Massiah, A. (2014). TEMPRANILLO is a regulator of juvenility in plants. *Sci. Rep.* 4:3704. doi: 10.1038/srep03704

Simpson, D. W., and Sharp, D. S. (1988). The inheritance of fruit yield and stolon production in everbearing strawberries. *Euphytica* 38, 65–74. doi: 10.1007/BF00024812

Song, Y. H., Shim, J. S., Kinmonth-Schultz, H. A., and Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. *Ann. Rev. Plant Biol.* 66, 441–464. doi: 10.1146/annurev-arplant-043014-115555

Sønsteby, A., and Heide, O. M. (2006). Dormancy relations and flowering of the strawberry cultivars Korona and Elsanta as influenced by photoperiod and temperature. *Sci. Horticult.* 110, 57–67. doi: 10.1016/j.scienta.2006.06.01

Sønsteby, A., and Nes, A. (1998). Short days and temperature effects on growth and flowering in strawberry (*Fragaria* × *ananassa* Duch.). *J. Horticult. Sci. Biotechnol.* 73, 730–736. doi: 10.1080/14620316.1998.11511040

Stewart, P. J., and Folta, K. M. (2010). A review of photoperiodic flowering research in strawberry (*Fragaria* spp.). *Crit. Rev. Plant Sci.* 29, 1–13. doi: 10.1080/07352680903436259

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

Tenreira, T., Pimenta Lange, M. J., Lange, T., Bres, C., Labadie, M., Monfort, A., et al. (2017). A Specific Gibberellin 20-oxidase dictates the flowering-runnering decision in diploid strawberry. *Plant Cell* 29, 2168–2182. doi: 10.1105/tpc.16.00949

Thompson, P., and Guttridge, C. (1959). Effect of gibberellic acid on the initiation of flowers and runners in the strawberry. *Nature* 184, BA72–BA73.

Valverde, F. (2011). CONSTANS and the evolutionary origin of photoperiodic timing of flowering. *J Exp. Bot.* 62, 2453–2463. doi: 10.1093/jxb/erq449

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., and Coupland, G. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003–1006. doi: 10.1126/science.1091761

van Engelen, F. A., Molthoff, J. W., Conner, A. J., Nap, J. P., Pereira, A., and Stiekema, W. J. (1995). pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgen. Res.* 4, 288–290. doi: 10.1007/BF01969123

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 September 2022
DOI 10.3389/fpls.2022.971506

OPEN ACCESS

EDITED BY
Shunquan Lin,
South China Agricultural University, China

REVIEWED BY
Mingjun Li,
Northwest A&F University, China
Qiong Lin,
Chinese Academy of Agricultural Sciences, China

*CORRESPONDENCE
Anping Guo
gap211@126.com
Xiumei Zhang
asiazhang1975@163.com

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 17 June 2022
ACCEPTED 23 August 2022
PUBLISHED 08 September 2022

CITATION
Gao Y, Yao Y, Chen X, Wu J, Wu Q, Liu S, Guo A and Zhang X (2022) Metabolomic and transcriptomic analyses reveal the mechanism of sweet-acidic taste formation during pineapple fruit development. *Front. Plant Sci.* 13:971506. doi: 10.3389/fpls.2022.971506

Metabolomic and transcriptomic analyses reveal the mechanism of sweet-acidic taste formation during pineapple fruit development

Yuyao Gao[1,2], Yanli Yao[2], Xin Chen[3], Jianyang Wu[4], Qingsong Wu[2], Shenghui Liu[2], Anping Guo[5]* and Xiumei Zhang[2]*

[1]College of Tropical Crops, Hainan University, Haikou, China, [2]Key Laboratory of Ministry of Agriculture for Tropical Fruit Biology, South Subtropical Crop Research Institute, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, China, [3]Taixing Institute of Agricultural Sciences, Taixing, China, [4]Department of Science Education, Zhanjiang Preschool Education College, Zhanjiang, China, [5]Sanya Research Institute, Chinese Academy of Tropical Agricultural Sciences, Sanya, China

Pineapple (*Ananas comosus* L.) is one of the most valuable subtropical fruit crop in the world. The sweet-acidic taste of the pineapple fruits is a major contributor to the characteristic of fruit quality, but its formation mechanism remains elusive. Here, targeted metabolomic and transcriptomic analyses were performed during the fruit developmental stages in two pineapple cultivars ("Comte de Paris" and "MD-2") to gain a global view of the metabolism and transport pathways involved in sugar and organic acid accumulation. Assessment of the levels of different sugar and acid components during fruit development revealed that the predominant sugar and organic acid in mature fruits of both cultivars was sucrose and citric acid, respectively. Weighted gene coexpression network analysis of metabolic phenotypes and gene expression profiling enabled the identification of 21 genes associated with sucrose accumulation and 19 genes associated with citric acid accumulation. The coordinated interaction of the 21 genes correlated with sucrose irreversible hydrolysis, resynthesis, and transport could be responsible for sucrose accumulation in pineapple fruit. In addition, citric acid accumulation might be controlled by the coordinated interaction of the pyruvate-to-acetyl-CoA-to-citrate pathway, gamma-aminobutyric acid pathway, and tonoplast proton pumps in pineapple. These results provide deep insights into the metabolic regulation of sweetness and acidity in pineapple.

KEYWORDS

***Ananas comosus*, fruit quality, sucrose, citric acid, metabolic genes, transporter genes**

Introduction

Among the tropical fruits, pineapple (*Ananas comosus* L.) has a production volume of about 25 million metric tons and ranks third in the world after banana and citrus (Ali et al., 2020; Ikram et al., 2020). The useful fruit is processed into value-added compounds because of its pleasant flavor and high abundance of biologically active substances such as vitamins and phenolic and carotenoid compounds (Léchaudel et al., 2018; Ali et al., 2020). China is an important pineapple-producing country, and "Comte de Paris," the dominant cultivar, accounts for more than 90% of the total pineapple cultivation area (Lu et al., 2014; Zhang et al., 2015). However, because of intensive cultivation focused on yield, commercially produced pineapple fruit seems to have lost its distinct sweet-acidic taste and fruity aroma much like other fruit crops (Klee and Tieman, 2018). Thus, understanding the quality-related metabolic characteristics and genetic basis in pineapple fruit is essential to performing molecular breeding targeting flavor-modulating genes that can contribute to producing a healthy, favorite fruit for consumers.

Sweetness is an important determinant of fruit organoleptic quality, and its intensity is determined by the content and composition of soluble sugars, such as sucrose (Suc), and hexoses, including glucose (Glc) and fructose (Fru) (Vimolmangkang et al., 2016; Klee and Tieman, 2018). Sugar accumulation in fruit is an intricate process that mobilizes the long-distance transport of photoassimilates by sieve element–companion cell complexes of the phloem to sink cells for unloading by symplastic and/or apoplastic pathways (Li et al., 2012; Vimolmangkang et al., 2016). Suc enters the storage parenchyma cells by Suc transporters (SUTs) or hexose transporters after hydrolysis by cell wall invertases (Li et al., 2012). In parenchyma cells, sugar metabolism in the cytoplasm comprises an elaborate system termed the Suc-Suc cycle involving various enzymes, including Suc synthase (SUSY), Suc phosphate synthase (SPS), hexokinase (HK), and phosphofructokinase (PFK) (Li et al., 2012; Aslam et al., 2019). In the vacuole, imported Suc can be irreversibly hydrolyzed to hexose by vacuolar acid invertase (VINV) (Vimolmangkang et al., 2016). In addition to sugar metabolism, sugar accumulation in fruits is regulated by transmembrane transporters. Recently, key sugar transporters contributing to the distribution of sugars in fruit have also been identified, such as the monosaccharide transporter-like (MST), sugars will eventually be exported transporter (SWEET), and the Glc exporter early response to dehydration like 6 (ERDL6) (Wen et al., 2022).

Acidity is also a crucial contributor to fruit taste quality, and its perception depends on organic acid content and composition (Etienne et al., 2013; Klee and Tieman, 2018). The predominant organic acids in most fruits are citrate and malate, which accumulate by complex processes, including synthesis, transport, and degradation or utilization (Zheng et al., 2020; Huang et al., 2021). In the tricarboxylic acid (TCA) cycle, important enzymes such as citrate synthase (CS), aconitase (ACO), malate dehydrogenase (MDH), and isocitrate dehydrogenase (ICDH) play an essential role in citrate and malate metabolism in fruits. In the cytosol, citrate can be degraded or utilized by the gamma-aminobutyric acid (GABA) or acetyl-CoA pathways that affect fruit acidity (Zheng et al., 2020). In addition to organic acid metabolism, acid accumulation is governed by the transport of organic acids from the cytosol to the vacuole. Recently, vacuolar transporters and ion channels and carriers have been reported to play a major role in controlling fruit acidity; for example, the tonoplast proton pumps such as vacuolar-type (V-ATPase, V-PPase) and p-type ATPases, and the channels of aluminum-activated malate transporters (ALMTs) (Huang et al., 2021). In addition to acidic acids and other organic acid components, chlorogenic acid is an important phenolic acid determining the antioxidant capacity and edible value of fruit crops (Chen et al., 2020).

Pineapple is categorized as a non-climacteric tropical fruit and experiences complex biochemical changes, such as variations in the metabolism of sugar, organic acids, and phenolic compounds during fruit development (Saradhuldhat and Paull, 2007; Zhang et al., 2012; Léchaudel et al., 2018). An assessment of sugars and organic acid levels in mature fruits of 26 pineapple cultivars by high-performance liquid chromatography showed that the predominant sugar was Suc, followed by Glc and Fru, and the prevailing acid was citric acid, accompanied by malic and quinic acids (Lu et al., 2014). Sugars and organic acids are not only important sources of fruit attributes, such as sweetness, acidity, and antioxidants, but also play a key role in regulating plant development, stress resistance, and yield as signal molecules. Thus, understanding sugar and organic acid accumulation is of great significance

Abbreviations: Suc, sucrose; Glc, glucose; Fru, fructose; PEP, phosphoenolpyruvate; GABA, gamma-aminobutyric acid; SUT, sucrose transporter; INV, acid invertase; CWINV, cell wall invertase; INH, invertase inhibitor; SUSY, sucrose synthase; SPS, sucrose phosphate synthase; HK, hexokinase; PFK, phosphofructokinase; VINV, vacuolar acid invertase; MST, monosaccharide transporter-like; SWEET, sugars will eventually be exported transporter; ERDL6, the glucose exporter early response to dehydration like 6; CS, citrate synthase; ACO, aconitase; MDH, malate dehydrogenase; ICDH, isocitrate dehydrogenase; ALMTs, aluminum-activated malate transporters; PFK, phosphofructokinase; PDH, pyruvate dehydrogenase; PK, pyruvate kinase; NADP-IDH, isocitrate dehydrogenase NADP; GAD, glutamate decarboxylase; H^+-ATPase, vacuolar-type and p-type ATPase; NaDCs, sodium-dependent dicarboxylate transporter; RNA-Seq, next-generation RNA sequencing; LC-MS/MS, metabolomics; PCA, principal component analysis; HCA, hierarchical cluster analysis; VIP, variable importance in projection; MRM, multiple reaction monitoring; WGCNA, weighted gene coexpression network analysis; OPLS-DA, orthogonal partial least squares discriminant analysis; QC, quality control; qRT-PCR, quantitative real-time PCR; FPKM, fragments per kilobase of transcript per million fragments mapped.

(Chen et al., 2020; Huang et al., 2021; Wen et al., 2022). As mentioned above, the metabolism and accumulation of sugars and organic acids are controlled by complex metabolic pathways and transmembrane transporters involved in multigene responses. Compared with other horticultural crops, such as apple and peach (Li et al., 2012; Zheng et al., 2020), the information on metabolic modulations and accumulation mechanisms related to sugars and organic acids in pineapple is limited. Although some studies have been reported on the changes in key enzymes and relevant biosynthesis genes linked to sugars and organic acids during pineapple fruit development (Saradhuldhat and Paull, 2007; Zhang et al., 2012, 2019; Wu et al., 2022), the metabolic pathways and genes controlling sugar and organic acid accumulation in pineapple fruit remain largely unknown.

The release of the high-quality pineapple genome database paves the way for clarifying the molecular mechanisms in pineapple based on genome-wide transcriptomic analysis (Ming et al., 2015; Wang L. L. et al., 2020). Moreover, with the advances in transcriptomic analysis approaches, a useful method such as weighted gene coexpression network analysis (WGCNA) has made it possible to detect modules of coexpressed genes and key genes responsible for important traits by correlating the transcriptome data with phenotypic data, as reported in watermelon and kiwifruit with different genotypes (Umer et al., 2020; Liao et al., 2021). In addition, although the metabolic phenotype is largely influenced by the environment and difficult to monitor accurately, a rapid and reliable metabolomics approach such as liquid chromatography with tandem mass spectrometry (LC-MS/MS) can detect a range of plant metabolites with high throughput and sensitivity from different biological samples (Hong et al., 2021). Recently, transcriptomics and metabolomics have been integrated to obtain the gene networks and regulatory mechanisms of quality formation in horticultural plants, such as watermelon (Umer et al., 2020) and kiwifruit (Wang et al., 2022). However, little research has been reported on combined RNA sequencing (RNA-Seq) and LC-MS/MS in the study of the molecular mechanisms controlling organic acid and sugar accumulation in pineapple.

In this study, LC-MS/MS and RNA-Seq analyses were used to investigate the primary metabolic dynamics and gene expression profiles of two pineapple cultivars in different fruit developmental stages, focusing on the metabolites and genes associated with sugars and organic acids contributing to fruit quality. Furthermore, WGCNA was performed to construct the regulatory networks and screen related genes controlling the accumulation of sugar and organic acids by the combination of these metabolic data and gene expression profiles. Overall, these data can provide important insights into the regulation of sugar and organic acid metabolism and establish a theoretical basis for flavor improvement in pineapple.

Results

Morphological and physiological characteristics of pineapple fruit during development

"Comte de Paris" (B) is an early mature cultivar widely cultivated in China, and "MD-2" (M) is a newly introduced middle mature cultivar. Changes in fruit weight, total soluble solids (TSS), total acids (TA), and TSS/TA ratio during different developmental stages of the two cultivars are shown in **Figure 1**. The fruit weight of both cultivars consistently increased during the rapid expansion stage [20–90 days after anthesis (DAA) for M, 20–70 DAA for B], and then maintained a limited variation during the mature stage (90–120 DAA for M, 70–100 DAA for B), with a gradual change in fruit color from green to yellow (**Figures 1A,B**). In common practice, fruit harvest of M is usually at 120 DAA and that of B is at 100 DAA, when more than two-thirds of the fruit has turned yellow and reached 80% maturity.

The total soluble solids (TSS) content of the two varieties increased slightly during the expansion stage, and sharply increased during the maturity stage, and then reached the maximum at harvest time (**Figure 1C**). The TSS content of M was lower than that of B during the 20–100 DAA, but there was no pronounced difference at harvest time. The total acids (TA) content of both cultivars gradually increased and peaked at 20 days before harvest, and then decreased slightly until harvest (**Figure 1D**). The TA content of M was higher than that of B during the 20–90 DAA, but there was no significant difference at harvest time. The TSS/TA ratio of M differed slightly from 20 to 80 DAA and then increased till 120 DAA, and the ratio in B differed slightly from 20 to 60 DAA, then increased till 100 DAA (**Figure 1E**). The differences in TSS, TA and TSS/TA ratio at different stages accounted for the differences in fruit sweet-acidic taste at different stages of the M and B cultivar.

To investigate the soluble sugars and organic acid-associated metabolic modulations throughout the development of pineapple fruit flesh, four developmental stages (40, 80, 100, and 120 DAA; represented by M1, M2, M3, and M4) for M and three developmental stages (40, 80, and 100 DAA, represented by B1, B2, and B3) for B were selected and used for transcriptome sequencing and metabolomic analysis.

Metabolic profiling of pineapple fruit during development

Targeted metabolomic analysis based on ultra-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (UPLC-ESI-MS/MS) was performed to investigate the primary metabolite (e.g., sugars and organic acids)

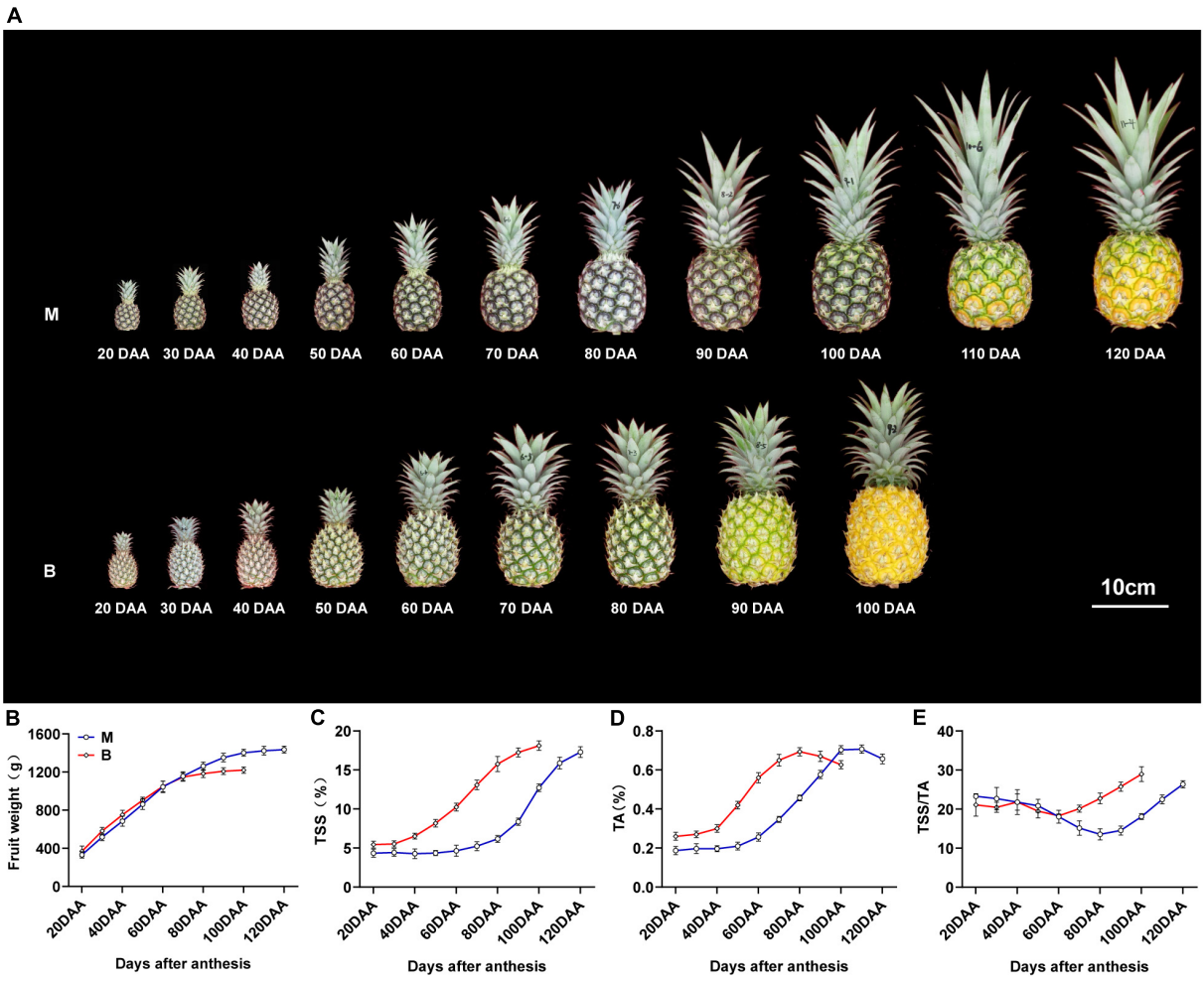


FIGURE 1
Changes in fruit morphology and physiological index and fruit weight during fruit development and maturity. **(A)** Pineapple cultivars M and B fruits at different developmental stages. **(B)** Fruit weight in the fruit at different developmental stages. **(C)** TSS in the fruit at different developmental stages. **(D)** TA in the fruit at different developmental stages. **(E)** TSS/TA in the fruit at different developmental stages. DAA, days after anthesis; Data were the mean ± standard error from three biological replicate assays.

dynamics of pineapple fruit during development. A total of 280 primary metabolites, including 56 amino acids and their derivatives, 81 phenolic acids, 31 nucleotides and their derivatives, 24 sugars and alcohols, 28 organic acids, 52 lipids, and 8 vitamins, were identified in our study (Supplementary Table 1). In the principal component analysis (PCA) diagram (Figure 2A), pineapple fruit samples were separated into distinct clusters based on different developmental stages of the two cultivars, indicating the significant differences in metabolite levels among pineapple fruit samples during fruit development. The PCA results showed that the first principal component (PC1) explained 41.16% of the total variance and distinguished samples based on different developmental stages (rapid expansion or mature stages); the second principal component (PC2) separated the two cultivars with a 14.01% variance contribution value. A larger value of PC1 revealed that different developmental stages mainly caused differences in metabolites in samples. In the heatmap (Figure 2B), two main clusters were obtained according to metabolic accumulation during fruit development in the two cultivars. Metabolites such as phenolic acids, amino acids and derivatives and lipids in cluster 1 mainly accumulated at the rapid expansion stages (M1, M2, and B1), while organic acids, saccharides and alcohols, and some phenolic acids in cluster 2 preferentially accumulated at the mature stages (M3, M4, B2, and B3) (Table 1), indicating that metabolite accumulation was stage-specific during the expansion and mature stages of fruit development in pineapple. Moreover, both PCA and hierarchical clustering analysis (HCA) confirmed that the three biological replicates in each group were clustered, indicating the high reliability of the data.

Differential metabolites were analyzed based on variable importance in projection (VIP) $\geq$ 1 and | log2 (fold change)

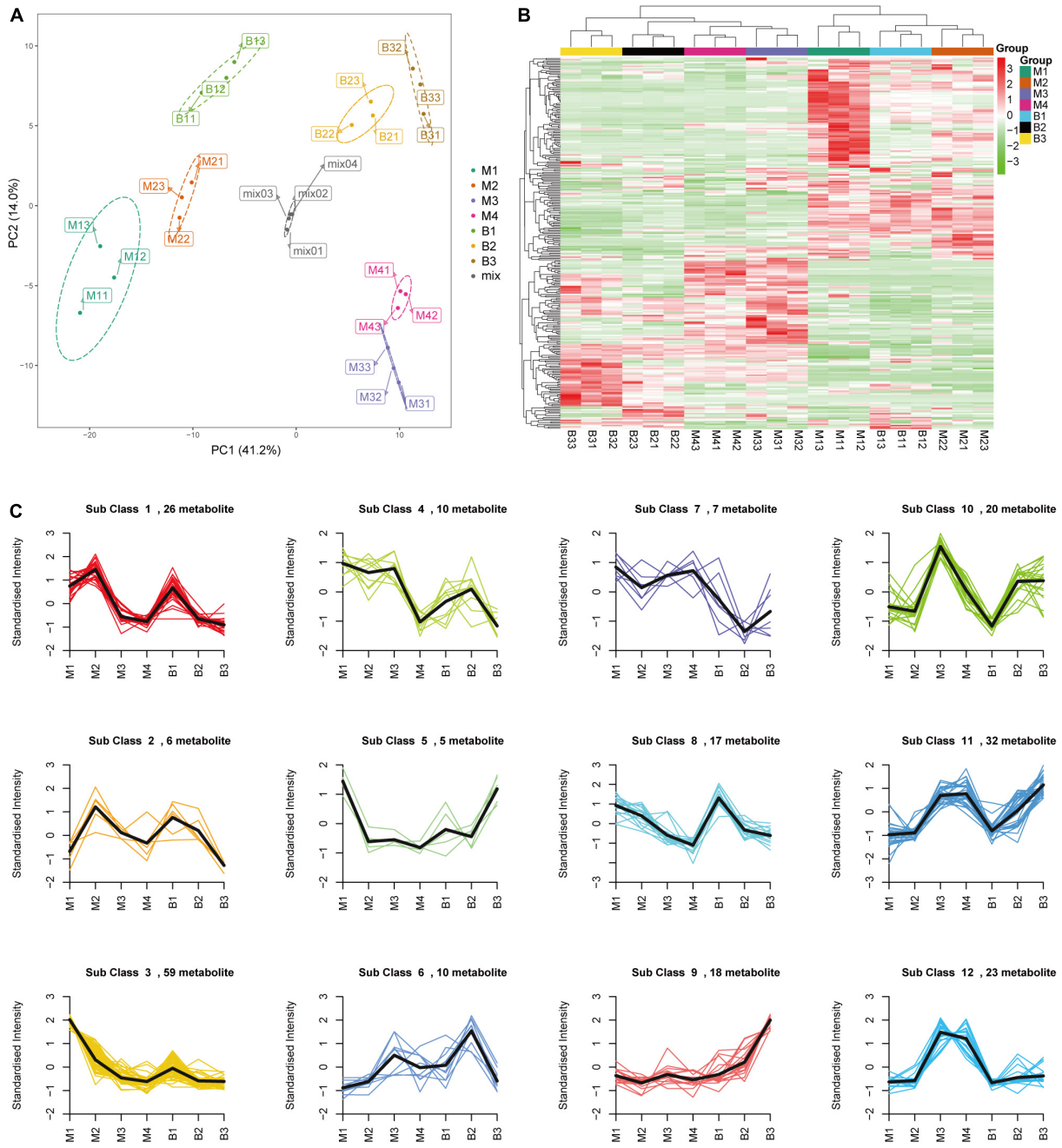


FIGURE 2
PCA **(A)**, HCA **(B)**, and K-means clustering **(C)** of the metabolites from the developing pineapple flesh. M1, M2, M3, and M4 represented fruit flesh samples for M at 40, 80, 100, and 120 DAA, respectively, B1, B2, and B3 represented fruit flesh samples for B at 40, 80, and 100 DAA, respectively. In the heatmap, each sample was represented by a single column, and each metabolite was visualized in a row. Red showed high abundance, and green exhibited relatively low metabolite abundance. 12 clusters (Sub class 1–12) of 233 differential metabolites were divided based on the dynamic changes of metabolites from different developmental stages for the two cultivars.

$|\ | \geq 1$. A total of 233 differential metabolites (18 sugars and alcohols, 24 organic acids, etc.) were identified at different developmental stages of the two cultivars (B1 vs. B2 and B3; B2 vs. B3; M1 vs. M2, M3, M4; M2 vs. M3, M4; M3 vs. M4) (Supplementary Table 2). In the orthogonal partial least squares discriminant analysis (OPLS-DA) results, the Q^2 values of the nine comparisons were higher than 0.87, indicating the satisfactory predictive capabilities of the models (Supplementary Figures 1A–I). To comprehensively understand the accumulation dynamics of 233 differential

TABLE 1 Distribution of 280 primary metabolites identified between cluster 1 and cluster 2 based on HCA.

Clusters	C1	C2	Total
Compounds	216	64	280
Phenolic acids	61	20	81
Lipids	45	7	52
Amino acids and derivatives	49	7	56
Organic acids	17	11	28
Saccharides and alcohols	14	10	24
Nucleotides and derivatives	24	7	31
Vitamin	6	2	8

metabolites with fruit development in both cultivars, K-means clustering analysis was applied to display 12 distinct subclasses (**Figure 2C** and **Table 2**). Among these subclasses, subclasses 5, 7, 9, 10, and 12 accounted for only a small proportion (31%) of metabolites and mainly featured phenolic acids and nucleotides and derivatives, and showed different change trend during the entire period of the fruit development between the two varieties. For subclasses 1, 2, 3, 4, 6, and 8, metabolites were mainly composed of lipids, amino acids and derivatives and phenolic acids, and displayed the significant decreasing during certain periods of the fruit development in both cultivars. However subclass 11 exhibited the significant upregulation of metabolites from the rapid expansion to mature stage in both cultivars, and mainly comprised of phenolic acids, sugars and organic acids, which were associated with the formation of fruit sweet-acidic taste. The change pattern of these metabolites in subclass 11 was similar to that of TSS and TA during the fruit development in both cultivars. Specifically, Suc and citric acid were found to be the predominant sugar and organic acid of many pineapple genotypes, respectively (Lu et al., 2014), and were concentrated in subclass 11. Additionally, chlorogenic acid was considered to be related to antioxidant-associated attributes of pineapple (Arampath and Dekker, 2019), were also grouped in subclass 11.

Variations among sugars and organic acids in pineapple fruit during development

Metabolites likely to be associated with sugar and organic acid metabolism and accumulation in pineapple, included four sugars (Sucrose, glucose 6-phosphate, glucose 1-phosphate, and glucose) an alcohol (sorbitol), six organic acids (citric acid, quinic acid, malic acid, fumaric acid, gamma-aminobutyric acid, and phosphoenolpyruvic acid), and a phenolic acid (chlorogenic acid) (**Figure 3**). For the four sugars and alcohol, Suc content was low during the rapid expansion stages and increased significantly during the mature stages in the two cultivars (**Figure 3A**), whereas Glc content was high during the rapid expansion stages but displayed a slight decreasing trend during the mature stages in the two cultivars (**Figure 3B**). Developmental changes in Glc-1-phosphate and Glc 6-phosphate content showed a decrease, whereas sorbitol content displayed no significant changes in the two cultivars (**Figures 3C–E**). Among the six organic acids, increasing trends were observed in the content of citric acid and quinic acid during fruit development in the two cultivars, and the citric acid level dramatically increased from the expansion to mature stages for both cultivar (**Figures 3J,K**). In contrast, the levels of malic acid, fumaric acid, GABA, and phosphoenolpyruvic acid displayed decreasing trends as the fruits moved toward maturity in the two cultivars, and malic acid levels exhibited significant downregulation from the expansion to mature stages for the two cultivars (**Figures 3G–I,L**). In addition, an increasing trend in the chlorogenic acid level was observed during fruit development in the two cultivars, and the chlorogenic acid content significantly increased from the expansion to mature stages in M cultivar (**Figure 3F**). Consistent with the previous results, Suc and citric acid contents were the highest in mature fruit of the soluble sugars and organic acids in both cultivars, respectively (**Table 3**) (Lu et al., 2014). Altogether, the changes in the Suc and citric acid contents during fruit development in both cultivars were consistent with that of the TA and TSS contents, respectively, indicating that Suc and citric acid could be crucial for the formation of fruit sweet-acidic taste of both cultivars.

Transcriptome profiling of pineapple fruit during development

To investigate the molecular basis underlying pineapple quality-associated metabolism changes, 21 cDNA libraries for different developmental stages of the two cultivars were constructed by RNA-Seq. The details of RNA-Seq data for each sample are shown in **Supplementary Table 3**. After filtering low-quality reads in each library, 7.48 Gb clean reads on average for each individual sample lines were generated with >89% Q30 base percentage. When clean reads were mapped to the pineapple genome, >80% of clean reads were matched (**Supplementary Table 3**). In addition, both PCA and HCA of the transcriptome confirmed a clear separation between the groups and a good correlation within the groups (**Figures 4A,B**), revealing the reliability of the transcriptome data.

Identification of coexpressed gene networks by weighted gene coexpression network analysis

To understand the regulatory mechanisms of soluble sugars and organic acids during pineapple fruit developmental stages, WGCNA was used to identify the coexpressed gene networks. After removing low- expressed genes from all identified genes (**Supplementary Table 4**), 3,723 genes with FPKM values

10.3389/fpls.2022.971506

TABLE 2 Distribution of the 233 differentially accumulated metabolites identified among different k-means clusters.

Subclass	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Total
Compounds	26	6	59	10	5	10	7	17	18	20	32	23	233
Phenolic acids	7	1	10	3	2	3	3	2	7	1	15	16	70
Lipids	7	2	23	2	1	4	0	0	0	5	0	2	46
Amino acids and derivatives	6	0	13	3	0	0	0	6	2	10	2	3	45
Organic acids	1	3	4	1	0	0	0	3	2	1	7	2	24
Saccharides and alcohols	2	0	4	0	0	0	2	1	1	0	8	0	18
Nucleotides and derivatives	2	0	4	1	2	2	1	4	6	3	0	0	25
Vitamin	1	0	1	0	0	1	1	1	0	0	0	0	5

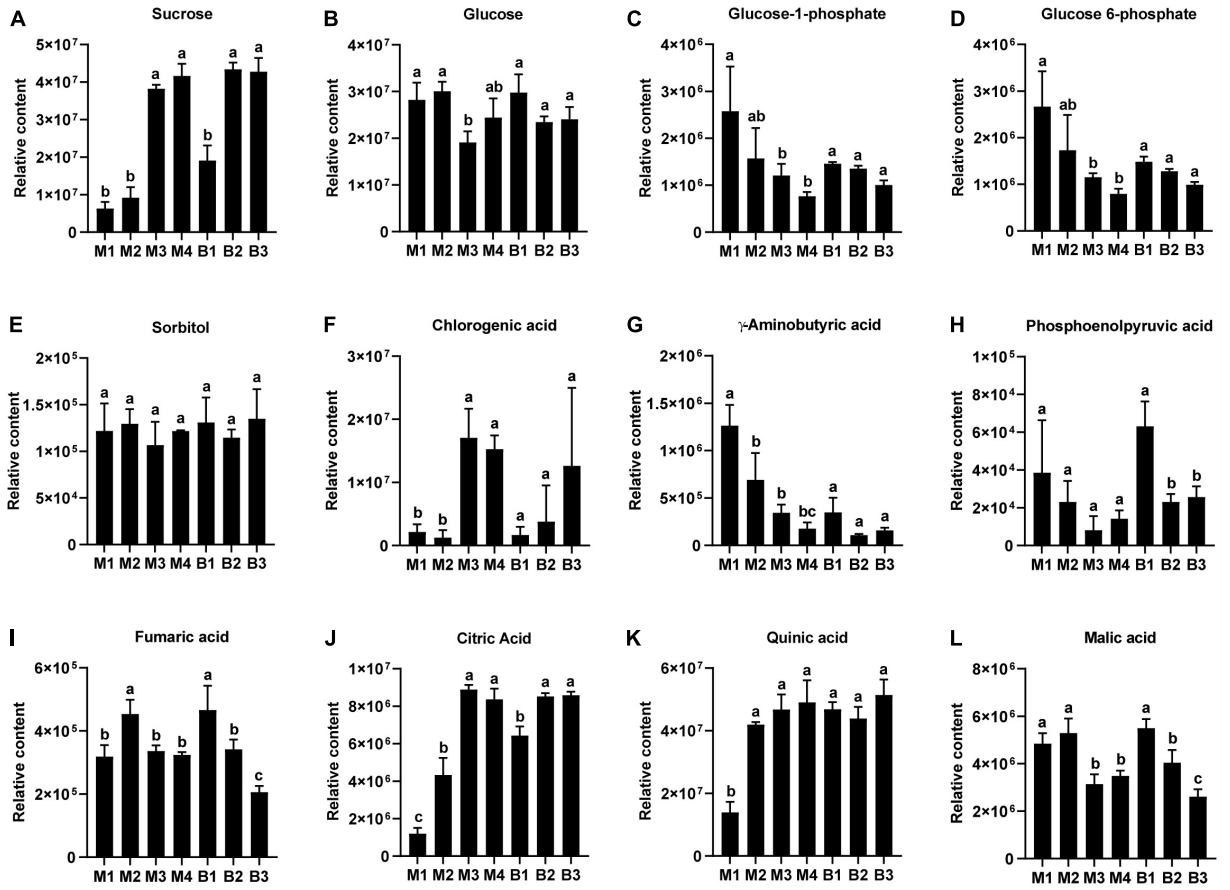


FIGURE 3
Changes in the relative contents of the main sugars and organic acids in pineapple fruits at four different stages (40, 80, 100, and 120 DAA) for M (M1, M2, M3, and M4) and three different stages (40, 80, and 100 DAA) for B (B1, B2 and B3). **(A)** Sucrose, **(B)** glucose, **(C)** glucose 1-phosphate, **(D)** glucose 6-phosphate, **(E)** sorbitol, **(F)** chlorogenic acid, **(G)** gamma-aminobutyric acid, **(H)** phosphoenolpyruvic acid, **(I)** fumaric acid, **(J)** citric acid, **(K)** quinic acid, **(L)** malic acid. For each cultivar, bars with the same lowercase letter (a, b, or c) indicate no significant differences ($p < 0.05$).

were entered into the WGCNA module. Here, Suc, Glc, quinic acid, citric acid, malic acid, and chlorogenic acid content at different stages of the two cultivars were taken as phenotypic data for the analysis of gene module–trait correlations. A total of 8 distinct modules marked with different colors were produced according to the similar coexpressed patterns of each gene (Figure 5A and Supplementary Table 5). Detailed information of the gene module–trait correlations is shown in Figure 5B. Interestingly, the turquoise and purple modules showed significant positive correlations with Suc, citric acid, and chlorogenic acid, which preferentially accumulated at mature stages (B2, B3, M3, and M4), whereas

10.3389/fpls.2022.971506

TABLE 3 Content of sugars and organic acids at different developmental stages in M and B cultivar (mg·g^{-1} FW) and the values shown are the mean ± SD.

	M1	M2	M3	M4	B1	B2	B3
Sucrose	14.22 ± 0.58	20.38 ± 0.55	68.9 ± 0.52	74.06 ± 0.66	25.92 ± 0.68	58.03 ± 0.98	63.23 ± 0.52
Fructose	5.41 ± 0.26	6.62 ± 0.24	11.51 ± 0.71	13.29 ± 0.47	12.61 ± 0.27	22.22 ± 0.44	28.14 ± 0.35
Glucose	6.67 ± 0.30	7.71 ± 0.31	14.02 ± 0.28	15.81 ± 0.50	14.44 ± 0.68	24.74 ± 0.40	31.50 ± 0.83
Citric acid	1.13 ± 0.16	2.29 ± 0.16	3.78 ± 0.15	3.35 ± 0.27	1.53 ± 0.12	3.37 ± 0.20	2.77 ± 0.32
Quinic acid	0.24 ± 0.02	0.67 ± 0.03	0.83 ± 0.02	0.93 ± 0.06	0.47 ± 0.03	0.65 ± 0.02	0.78 ± 0.02
Malic acid	1.57 ± 0.05	1.69 ± 0.03	1.15 ± 0.02	1.04 ± 0.04	1.64 ± 0.04	1.12 ± 0.02	0.85 ± 0.02

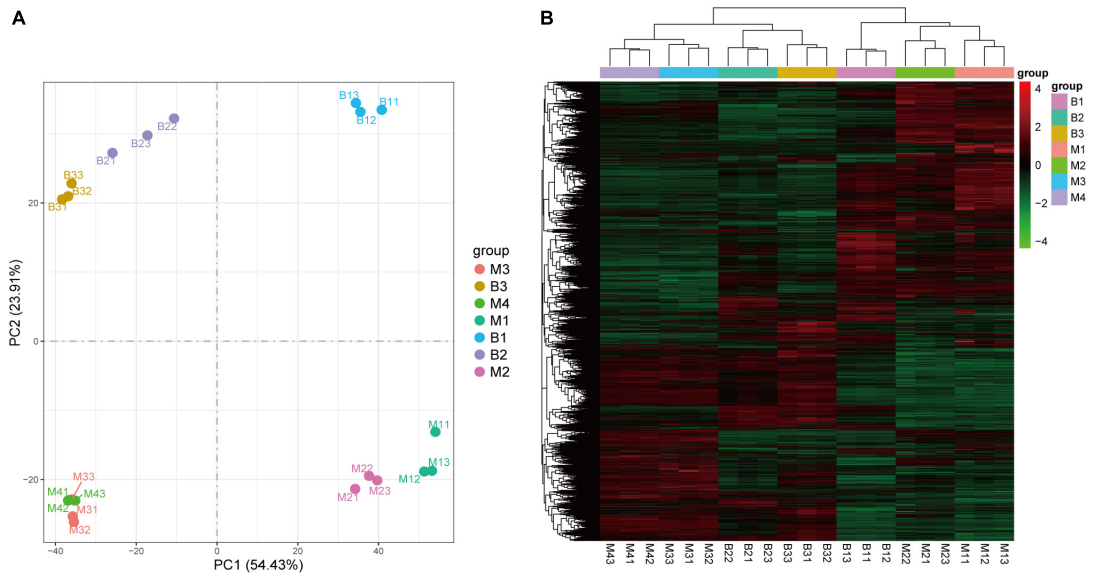


FIGURE 4
PCA **(A)** and HCA **(B)** of genes in pineapple fruits at different developmental stages of M and B. M1, M2, M3, and M4 represented fruit flesh samples for M at 40, 80, 100, and 120 DAA, respectively, B1, B2, and B3 represented fruit flesh samples for B at 40, 80, and 100 DAA, respectively.

the brown and cyan modules presented significant negative correlations with these metabolites mainly accumulating at the mature stages. In contrast, the turquoise and purple modules showed negative correlations with Glc and malic acid, which mainly accumulated at the rapid expansion stages (B1, M1, and M2), whereas the brown and cyan modules showed significant positive correlations with those metabolites mainly accumulating at the rapid expansion stages. Thus, the genes in the four modules might play a vital role in modulating soluble sugar and organic acid metabolism in the pineapple fruit developmental stages.

Quantitative real-time PCR verification of gene expression

To verify the RNA-Seq data, 15 genes linked to sugar and organic acid accumulation in pineapple flesh were subjected to qRT-PCR analysis (Figure 6). qRT-PCR demonstrated that the expression trends of the genes were consistent with the RNA-Seq analysis results (Figure 6A). Correlation analysis ($R^2 = 0.79$, $P < 0.0001$) between RNA-Seq and qRT-PCR data indicated the reliability of the RNA-Seq data (Figure 6B).

Genes related to sugar metabolism and accumulation in pineapple fruit

Suc, the main contributor to pineapple sweetness (Zhang et al., 2012, 2019), was the predominant soluble sugar of mature fruits in both cultivars. Interestingly, Suc accumulation was strongly negatively associated with the brown ($r^2 = -0.92$) and cyan ($r^2 = -0.91$) modules and highly positively correlated with the turquoise ($r^2 = 0.73$) and purple ($r^2 = 0.76$) modules. Finally, 21 candidate genes linked to sugar metabolism and sugar transport in pineapple (Supplementary Table 6 and Figure 7). Acid invertases (INVs), including cell wall invertase (CWINV) and vacuolar acid invertase (VINV), participate in the

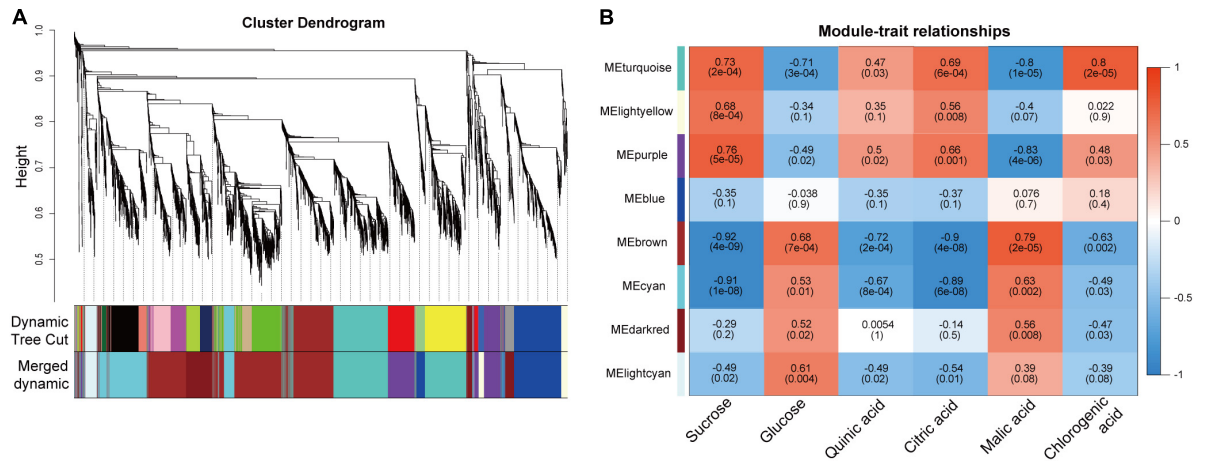


FIGURE 5
Result of gene coexpression modules associated with sugars and organic acids in pineapple fruits at different developmental stages of M and B. **(A)** Clustering dendrogram presenting 11 modules of coexpressed genes based on WGCNA. **(B)** Correlations coefficient and significance between modules and sucrose, glucose, quinic acid, citric acid, malic acid, chlorogenic acid, where each grid contained the corresponding correlation and *p*-value.

irreversible hydrolysis of Suc in the apoplast or vacuole. For Suc hydrolysis, two CWINV genes (*Aco023030*, *Aco023036*) in the turquoise and purple modules and a VINV gene (*Aco017533*) from the cyan module in pineapple were first identified. The expression levels of the two CWINV genes were positively correlated with Suc content, whereas the VINV gene expression level was strongly negatively correlated with Suc content during the developmental stages. Invertase inhibitor (INH), an inhibitory protein binding to the active site (Suc binding site) of INV, includes cell wall INHs (CWINHs) and vacuolar INHs (VINHs) based on their subcellular location (Wang et al., 2020; Kawaguchi et al., 2021). Here, of the four INH genes identified, the expression levels of *Aco018410* from the cyan module and *Aco019529* and *Aco014073* from the brown module were strongly negatively correlated with Suc content, and the *Aco013219* expression level from the turquoise module was positively correlated with Suc content during fruit development.

Within the Suc–Suc cycle, three Suc synthase (SUSY) genes (*Aco009508*, *Aco026490*, *Aco009217*) in the brown and cyan modules were strongly negatively correlated with Suc content during fruit development. In addition, one Suc phosphate synthase (SPS) gene (*Aco017533*) and one hexokinase (HK) gene (*Aco000407*) were found in the turquoise module, and their expression levels were positively correlated with Suc content during fruit development. Of the two phosphofructokinase (PFK) genes identified, *Aco000731* in the cyan module was strongly negatively correlated with Suc content, whereas *Aco009949* was positively correlated with Suc content throughout fruit development. In addition, seven sugar transporter genes, including two sucrose transporter (SUT) genes, four sugars will eventually be exported transporter (SWEET) genes, and one the Glc exporter early response to dehydration like 6 (ERDL6) gene, were examined. The expression level of one SUT gene (*Aco000269*) from the turquoise module was positively correlated with Suc content, whereas that of the other SUT gene (*Aco004135*) from the brown module was negatively correlated with Suc content during fruit development. Among the four SWEET genes, three SWEET genes (*Aco005793*, *Aco004463*, *Aco006347*) that negatively correlated with Suc were present in the brown module, and one SWEET gene (*Aco016508*) that positively correlated with Suc was found in the turquoise module. An ERDL6 gene (*Aco018780*) that positively correlated with Suc was found in the turquoise module.

Genes correlated with organic acid metabolism and accumulation in pineapple fruit

Citric acid, contributing to the sour taste of pineapple (Saradhuldhat and Paull, 2007; Lu et al., 2014), was also the major organic acid of mature fruits in the two cultivars. Altogether, 19 candidate genes were first identified as being related to organic acid metabolism and transport in pineapple (**Supplementary Table 7** and **Figure 8**). For citrate synthesis, nine genes, including four pyruvate kinase (PK) genes (*Aco010310*, *Aco007762*, *Aco026402*, *Aco006253*), one pyruvate dehydrogenase (PDH) gene (*Aco007928*), two citrate synthase (CS) genes (*Aco024619*, *Aco004295*), one aconitase (ACO) gene (*Aco014852*), and one isocitrate dehydrogenase NADP (IDH-NADP) gene (*Aco005500*), were identified. These nine genes were present in the turquoise or purple module, and their expression levels were highly positively correlated with citrate

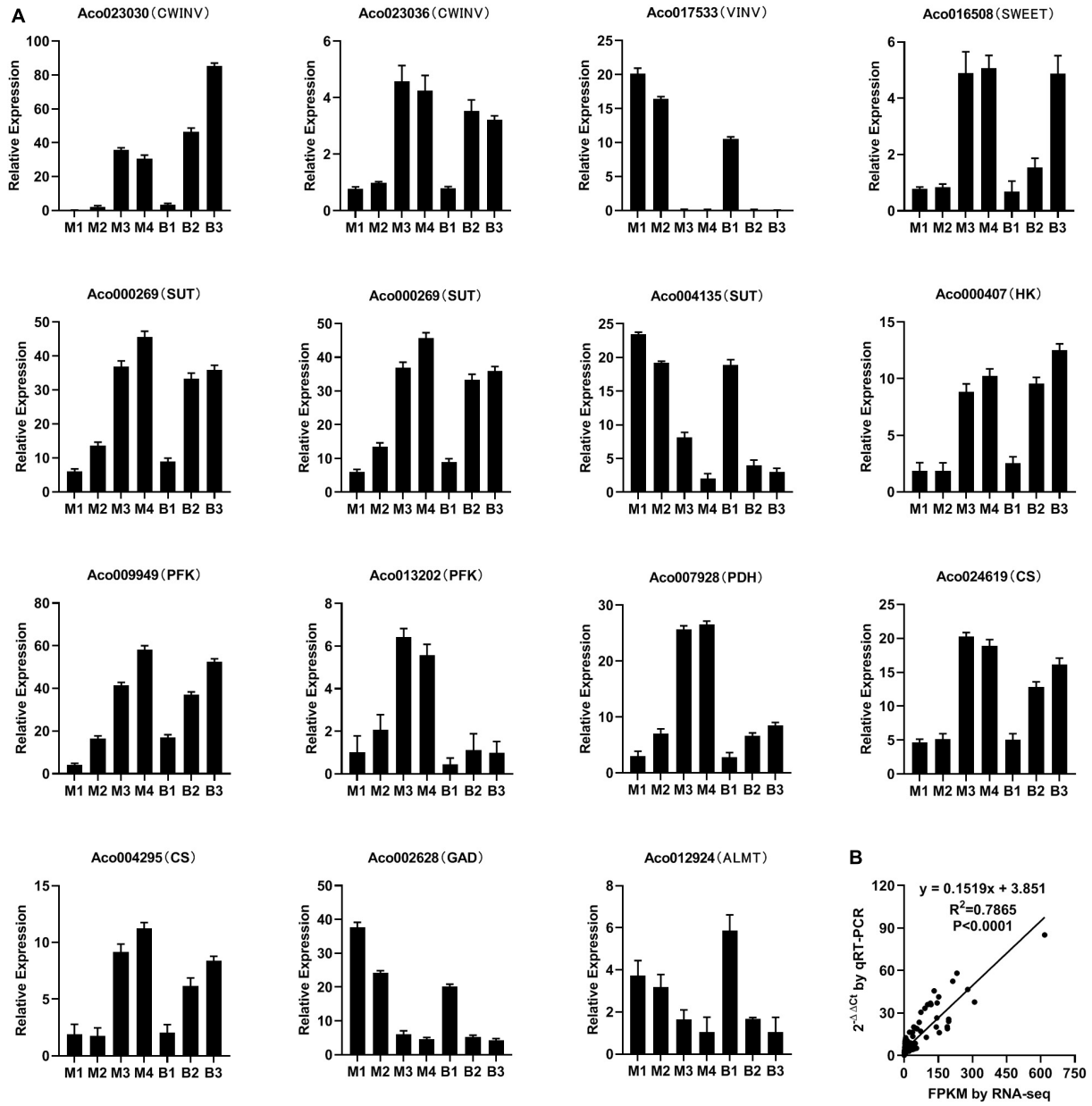


FIGURE 6
Relative expression of 15 genes involved in sugar and organic acid metabolism and transport during different developmental stages of M and B by qRT-PCR **(A)** and correlation analysis between RNA-Seq and qRT-PCR data **(B)**.

content (r^2 = 0.73 or r^2 = 0.76). For citrate degradation, 16 glutamate decarboxylase (GAD) genes were present in the transcriptome data, but only two GAD genes *Aco002628* and *Aco031209* were present in the brown and cyan modules, respectively, and were strongly negatively associated with citrate content (r^2 = –0.90, r^2 = –0.89). For citrate transport, four genes from the turquoise or purple module were identified as H^+-ATPase genes (*Aco002130*, *Aco024110*, *Aco022750*, *Aco009567*) involved in the diffusion of citrate into the vacuole. The expression levels of these four genes displayed positive correlations (r^2 = 0.73 or r^2 = 0.76) with citrate content. For malate metabolism and transport, two malate dehydrogenase (MDH) genes (*Aco014690*, *Aco004996*) in the purple module had a highly negative correlation (r^2 = –0.83) with malate content, whereas one aluminum-ALMT gene (*Aco012924*) and one sodium-dependent dicarboxylate transporter (NaDC) gene (*Aco000795*) in the brown module were highly positively correlated (r^2 = 0.79) with malate content.

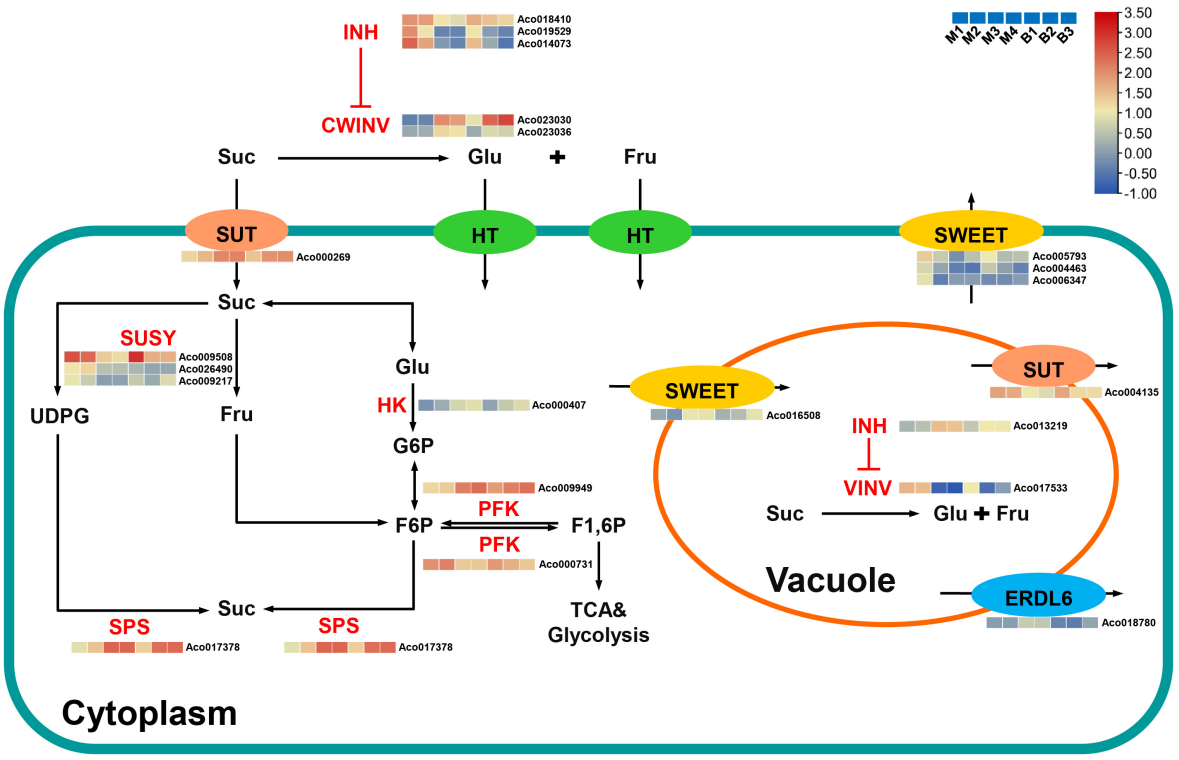


FIGURE 7
Regulatory model and expression levels of the genes related to sucrose metabolism and accumulation in pineapple fruits. From blue to red in the heatmap indicates the expression levels of the genes ranging from low to high in two pineapple cultivars throughout fruit development. M1, M2, M3, and M4 represented fruit flesh samples for M at 40, 80, 100, and 120 DAA, respectively, and B1, B2, and B3 represented fruit flesh samples for B at 40, 80, and 100 DAA, respectively. CWINV, cell wall invertase; CWINH, cell-wall invertase inhibitor; VINV, vacuolar acid invertase; VINH, vacuolar acid invertase inhibitor; SUSY, sucrose synthase; SPS, sucrose phosphate synthase; HK, hexokinase; PFK, phosphofructokinase; SUT, sucrose transporters; SWEET, sugars will eventually be exported transporter; ERDL6, glucose exporter early response to dehydration like 6; Suc, sucrose; Glc, glucose; Fru, fructose; UDPG, UDP-glucose; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; F1, 6P, fructose-1, 6-phosphate.

Discussion

The metabolism of sugars and organic acids is essential for fruit development, and their accumulation in fleshy fruits is central to the taste and quality of fruits (Chen et al., 2020; Huang et al., 2021; Wen et al., 2022). Primary metabolites, such as Suc, citrate, and malate, are important metabolites that not only provide energy and substrates for respiratory processes but are also a major source of fleshy fruit sweetness and acidity (Igamberdiev and Eprintsev, 2016; Wen et al., 2022). However, the description of the pineapple fruit metabolic spectrum mainly focuses on the mature stage or volatile metabolites (Steingass et al., 2015; Hong et al., 2021), and the global metabolic spectrum of different fruit developmental stages based on UPLC-ESI-MS/MS has not been identified. In this study, 280 primary metabolites were detected and annotated, and 233 differential metabolites were found in the two cultivars at different developmental stages. Metabolite accumulation in pineapple fruit was stage-specific (expansion and mature stages), which is consistent with the result in kiwifruit (Wang et al., 2022). Furthermore, by comprehensive transcriptomic and metabolomic analyses during fruit development, key metabolic pathways and complex genetic factors regulating sugar and organic acid accumulation were identified. In addition, the generated dataset could identify other pathways and regulatory genes related to flavor metabolism. Altogether, our finding will not only enhance the understanding of the complex regulatory mechanisms of sweet-acidic taste formation in pineapple but also provide valuable reference for the high-quality genetic improvement of pineapple.

Reasons for sucrose accumulation during pineapple fruit development

In this study, Suc was the predominant sugar in mature fruits and presented more remarkable changes during fruit development compared with the other sugars contributing to sweetness such as hexose, which was similar to previous

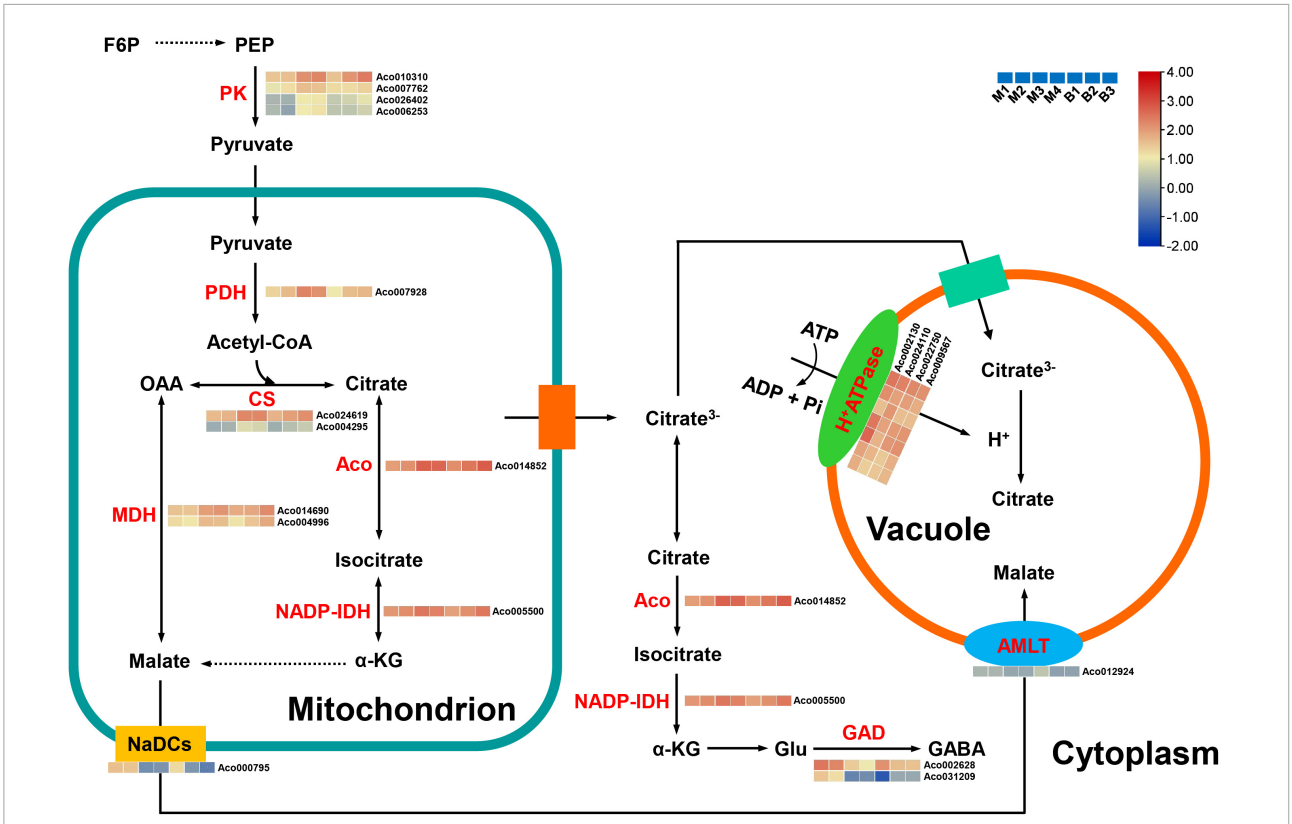


FIGURE 8
Regulatory model and expression levels of the genes related to citric acid and malic acid metabolism and accumulation in pineapple fruits. From blue to red in the heatmap indicates the expression levels of the genes ranging from low to high in two pineapple cultivars throughout fruit development. M1, M2, M3, and M4 represented fruit flesh samples for M at 40, 80, 100, and 120 DAA, respectively, and B1, B2, and B3 represented fruit flesh samples for B at 40, 80, and 100 DAA, respectively. PK, pyruvate kinase; PDH, pyruvate dehydrogenase; CS, citrate synthase, ACO, aconitase; MDH, malate dehydrogenase; NADP-IDH, isocitrate dehydrogenase NADP; GAD, glutamate decarboxylase; ALMT, aluminum-activated malate transporter; NaDCs, sodium-dependent dicarboxylate transporter. PEP, phosphoenolpyruvate; a-KG, a-ketoglutarate; Glu, glutamate; GABA, γ-aminobutyric acid.

studies (Zhang et al., 2012, 2019). As in apple and kiwifruit (Li et al., 2012; Vimolmangkang et al., 2016), the dataset in this study distinctly showed that Suc metabolism was highly developmentally modulated in pineapple. At the fruit expansion stage of the two cultivars, high expression levels of six genes on sugar metabolism were found, including three SUSY genes and one PFK genes (*Aco000731*), which may result in rapid metabolism of the imported Suc to supply the energy and intermediates for expansion and growth (Li et al., 2012), and one VINV gene, which could hydrolyze unmetabolized Suc to hexose for vacuolar accumulation. Thus, Suc content was low. As the pineapple fruit transitions to the mature stage, the expression levels of these six genes notably decreased, whereas the expression levels of two CWINV genes, one SPS gene, one HK gene, one PFK genes (*Aco009949*), one SWEET gene (*Aco016508*), one SUT gene (*Aco000269*), and one ERDL6 gene significantly increased. This process might be involved in CWINV-mediated Suc apoplast hydrolysis, SPS-mediated Suc cytoplasmic resynthesis, and SUT, SWEET, and ERDL6-mediated sugar transport. Thus, Suc content was high. Collectively, Suc accumulation in pineapple fruit is the result of coordinated actions of multiple genes.

CWINV catalyze the irreversible hydrolysis of Suc into glucose and fructose in the apoplast, determining the ratio of Suc: (glucose + fructose) in the apoplast of fruits during ripening (Walker et al., 2021). High CWINV activity contributed to the maintenance of a high intracellular-to-extracellular Suc ratio and the production of a negative water potential, leading to the continuous export of Suc into the apoplast (Minami et al., 2021). Apoplastic hexose could be directly transported into the cytosol to participate in Suc resynthesis (Miron et al., 2002). In this study, we identified two CWINV genes, which were highly expressed in fruits during mature stages, and displayed similar expression patterns to that of the Suc during fruit development, indicating that Suc hydrolysis in the apoplast mediated by CWINVs might be a critical step for Suc accumulation in pineapple, as reported in strawberry fruit (Yuan et al., 2021). CWIN activity could be regulated by invertase inhibitors (INHs) at the posttranslational level (Ru et al., 2020). Recently, the

knockout of *SlCWINH1* by CRISPR/Cas9 significantly increased the contents of Suc and hexose in genome-edited tomato line 193-3, possibly caused by higher CWINV activity in the apoplast of tomato fruits (Kawaguchi et al., 2021). In the present study, significant negative correlations of three INH genes with Suc implied that they might act as essential regulators of apoplastic Suc hydrolysis to control Suc accumulation in pineapple fruit. Unlike CWINVs, VINV mainly catalyze the irreversible hydrolysis of Suc into glucose and fructose and helps to maintain cell osmotic potential in vacuole (Ruan et al., 2010). The presence of VINV could contribute to increasing turgor pressure of expanding cells in fruits before ripening (Walker et al., 2021). The negative correlation of the VINV gene (*Aco017533*) with Suc content in this study showed that Suc was hydrolyzed into hexose by VINV in the vacuole at the expansion stages with less accumulation of Suc, and less likely to be hydrolyzed into hexose by VINV in the vacuole at the mature stages with Suc accumulation. Studies also show that post-translational limitation of VIN by INH contributes to control Suc content in fruits such as peach (Wang et al., 2020) and pear (Ma et al., 2020). Interestingly, the positive correlation of the INH gene (*Aco013219*) with Suc content of pineapple fruit suggested an interaction between Aco013219 and VINV, thereby participating in regulating pineapple fruit development and Suc accumulation. Hence, further studies are necessary to explore the interaction between INVs and INHs and their function in pineapple sugar accumulation.

Sugar transporters, the main members of the channels regulating sugar influx or efflux, significantly affect fruit sugar accumulation (Vimolmangkang et al., 2016; Aslam et al., 2019; Wen et al., 2022). Downregulation of *PpSUT4* expression in mature fruits showed the vacuolar efflux of Suc in peach (Zanon et al., 2015a,b). The *MdSUT4.1* expression level was negatively correlated with fruit sugar content, and its overexpression in strawberry and apple callus reduced sugar levels, suggesting that it was involved in remobilizing vacuolar sugar (Peng et al., 2020). In melon fruit, the high expression of the *CmSUT3* gene may be involved in the unloading of Suc in the apoplast (Wen et al., 2022). In the present study, the SUT gene (*Aco000269*) expression level showing a positive correlation with Suc content suggested that it was involved in apoplast unloading of Suc, whereas the SUT gene (*Aco004135*) expression level displaying a negative correlation with Suc content suggested that it was involved in Suc efflux from the vacuole membrane to the cytoplasm. Furthermore, the positive correlation between the SWEET gene (*Aco016508*) expression level and Suc content in this study showed that it might mediate Suc release into the apoplast and contribute to Suc accumulation in pineapple fruit based on a study in *Citrus* (Feng et al., 2021). The expression levels of the other three SWEET genes were negatively correlated with Suc content, which was consistent with previous studies (Guo et al., 2018). Thus, these three SWEET genes might transport Suc across the plasma membrane and be involved in Suc unloading as described in white pear (Li et al., 2017) or tomato (Zhang et al., 2021). In addition, although the hexose transporter MST1 gene has been identified in pineapple, its expression level at different fruit developmental stages remains unknown (Antony et al., 2008). ERDL6 is mainly responsible for transferring Glc from the vacuole to the cytosol. In apple, the expression levels of multiple *MdERDL6* family members were strongly positively related to Suc and Fru concentrations, and MdERDL6 mediated the Glc efflux to the cytoplasm, resulting in the accumulation of Suc and Fru in vacuolar by upregulating *tonoplast sugar transporter* expression (Zhu et al., 2021). Here, an *ERDL6* similar to *MST1* was identified, and its expression was positively correlated with Suc content during fruit development. This result further demonstrates that ERDL6 could be involved in regulating Suc accumulation in pineapple fruit. However, other hexose transporters, such as tonoplast sugar transporters and vacuolar Glc transporter, which are significantly related to sugar content of other fruits (Zhu et al., 2021; Wen et al., 2022), were not identified in the key modules. Nevertheless, we noticed that expression pattern of *ERDL6* (*Aco018780*) was similar to that of the *SWEET* (*Aco016508*), and SWEETs are regarded as having the characteristics of Suc transportation in the vacuolar membrane (Wen et al., 2022). It was speculated that, ERDL6-mediated Glc transportation could enhance the Suc accumulation in the vacuole of pineapple fruit mediated by the coordination of *ERDL6* (*Aco018780*) and *SWEET* (*Aco016508*).

Reasons for organic acid accumulation during pineapple fruit development

In the present study, early accumulation of malate was observed during the rapid expansion stages, and citrate content notably increased, whereas malate content significantly decreased as the fruits matured. In total, citric acid content were the highest in mature fruit of the organic acids in both cultivars. Thus, the citrate levels were also responsible for fruit acidity in both cultivars, which was consistent with previous research on pineapple (Lu et al., 2014; Wen et al., 2022). Owing to the vital role of organic acids in the regulation of osmotic pressure, stress resistance, and fruit quality (Huang et al., 2021), identifying the key genes and metabolic pathways regulating organic acid accumulation in pineapple is essential.

Citrate accumulation in fruit is controlled by citrate synthesis (Umer et al., 2020). In the last step of glycolysis, PK catalyzes the conversion of phosphoenolpyruvate (PEP) to pyruvate, and much pyruvate in the cytosol is transported to the mitochondria; subsequently, the PDH complex (composed of PDH, dihydrolipoyl transacetylase, and dihydrolipoyl dehydrogenase) converts pyruvate to acetyl-CoA, which determines the CS reaction rate in the TCA cycle (McCommis and Finck, 2015; Zheng et al., 2020). In peach fruit, the downregulation of PDH kinase gene (*PDK*) expression

inhibited PDH and the upregulation of *PK* expression in high-acid cultivars promoted citrate accumulation by the pyruvate-to-acetyl-CoA-to-citrate pathway (Zheng et al., 2020). A previous study in pineapple fruit revealed that the changes in CS and ACO activities coincided with citrate content during fruit development, and the decrease in CS activity and the increase in ACO activity were responsible for the reduction in the acidity in the low-acid clone compared with the high-acid clone (Saradhuldhat and Paull, 2007). In this study, four PK genes, one PDH gene, two CS genes, one Aco gene, and one IDH-NADP gene were positively correlated with citrate content. Thus, these nine genes are likely to play an important role in citrate synthesis during pineapple fruit development by the pyruvate-to-acetyl-CoA-to-citrate pathway.

Citrate degradation by the GABA pathway in the cytosol is also a key player in fruit acidity (Batista-Silva et al., 2018). GADs catalyze the decarboxylation of glutamate to GABA and are related to citrate accumulation, and the higher expression of GAD genes contribute to the decrease in fruit acidity at the late stage of development (Liu et al., 2014; Feng et al., 2021). Low-acid cultivars exhibited citrate degradation attributed to the upregulation of GAD gene expression in peach (Zheng et al., 2020). In this study, two GAD genes were notably negatively correlated with citrate content, suggesting that these two genes are crucial for citrate degradation by the GABA pathway during pineapple fruit development. Moreover, tonoplast proton pumps such as vacuolar-type and p-type ATPases play major roles in driving the facilitated diffusion of citrate into the vacuole (Huang et al., 2021). For instance, *CitPH1* and *CitPH5* encoding P-ATPase were expressed in acid varieties, whereas their expression was significantly reduced in acidless varieties, regulating hyper-acidification in citrus fruits (Strazzer et al., 2019). Another *CsPH8* gene in citrus has also been identified to play a role in determining citrate content or acidity (Shi et al., 2015, 2019). In the present study, one P-type ATPase gene and three H^+-ATPase genes were identified, and their expression levels were positively related with citrate content, suggesting that they might play central roles in citrate accumulation by driving citrate into the vacuole in pineapple. Besides, malate-related genes were identified in this study. Here, two MDH genes were positively correlated with malate content, whereas two malate-related transporter genes, including *NaDC* and *ALMT*, were significantly negatively correlated with malate content. This result showed that degradation and transport might regulate the malate content of pineapple.

Materials and methods

Plant materials

The conventional dominant cultivar "Comte de Paris" (B) belonging to "Queen," and excellent hybrid cultivar "MD-2" (M) belonging to "Smooth Cayenne," were used as experimental materials, which planted in the pineapple resource bank of South Subtropical Crops Research Institute (E 110°16′, N 21°10′, Zhanjiang, China) in 2019. All pineapple plants were uniformly managed during fertilization, irrigation, and disease and pest control. At the full florescence period, 150 pineapple plants per cultivar with normal growth and consistent flowering period were selected and tagged. The fruits per cultivar were randomly sampled at 10 a.m. every 10 days until harvest from 20 DAA. M at 11 developmental stages (20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120 DAA) and B at nine developmental stages (20, 30, 40, 50, 60, 70, 80, 90, 100 DAA) were collected (**Figure 1A**). The fruit weight, TSS and TA were determined throughout fruit development. Flesh from M at four developmental stages (40, 80, 100, 120 DAA) and B at three developmental stages (40, 80, 100 DAA) were collected according to the methods described in a previous study (Zhang et al., 2012) and immediately frozen in liquid nitrogen and stored at −80°C for further metabolic, transcriptomic, and qRT-PCR analyses. Measurements were made at each developmental stage with three biological replicates and nine fruits in each replicate.

Total soluble solids, total acids, sugars and organic acids measurement

The indicators of fruit quality mainly include total soluble solids (TSS), total acids (TA), TSS/TA ratio, sugars (glucose, fructose and sucrose) and organic acids (malic, citric and quinic acid). TSS (Brix%) was determined by using a hand-held refractometer (ATC-20E, Atago, Tokyo, Japan). Contents of TA, sugars and organic acids were determined by referring to our previous methods reported (Lu et al., 2014).

Targeted metabolomics analysis

The extraction, determination, and analysis of metabolites in the flesh samples were performed as previously described (Chen et al., 2013; Zou et al., 2020; Hong et al., 2021). In brief, the frozen fruit flesh samples were weighed, smashed, and extracted in 70% methanol solution for overnight at 4°C. Then, flesh extracts were centrifuged and filtered, and the resulting flesh samples were assessed by UPLC-ESI-MS/MS. During analysis, each of the 10 detected samples included a quality control sample using mixed flesh extracts to monitor repeatability and stability.

All primary metabolites were identified based on MetWare database[1] and quantified by multiple reaction monitoring. PCA, HCA, and OPLS-DA were performed in R to study differences

1 www.metware.cn

and reliability of metabolites from the 21 flesh samples. VIP in the OPLS-DA model was used to identify the differential metabolites (VIP $\geq$ 1 and | log2 (fold change) | $\geq$ 1). After standardized and centralized treatment, K-means analysis of differential metabolites was performed to study the relative content changes of metabolites in different samples.

Transcriptomics analysis

Total RNA was extracted from 21 frozen samples, and 21 cDNA libraries were constructed and sequenced on the Illumina Hiseq2000 platform with the generation of 150 bp paired-end reads as previously described (Umer et al., 2020). After processing the original and low-quality data (raw reads), clean reads were acquired and aligned to the pineapple reference genome[2] (Xu et al., 2018). The transcripts were annotated according to this pineapple genome database or the STRING database. The gene expression level was determined using the FPKM values based on gene lengths and read counts that were mapped to genes. In addition, both PCA and HCA were performed to evaluate differences between groups and the repetition of samples within groups.

Weighted gene coexpression network analysis

Coexpression network modules were constructed using the WGCNA package in R (Langfelder and Horvath, 2008; Umer et al., 2020). After introducing all genes into the WGCNA package, 3,723 genes with coefficient of variation (CV) > 0.5 and CV < standard deviation (sd) were screened to obtain the coexpression modules with the power of 6, minModuleSize of 30, and mergeCutHeight of 0.25 (Dai et al., 2021). After calculating the eigengene value of each module, the metabolic phenotype data was introduced, and the associations between phenotypes and gene modules during fruit development were obtained.

Verification of next-generation RNA sequencing data

The expression levels of the genes associated with sugar and organic acid accumulation were measured by qRT-PCR on a LightCycler480 II System (Roche, Switzerland) using primers listed in **Supplementary Table 8**. After the isolation of RNA of the 21 pineapple flesh samples, cDNA lines were generated using a BioTeke (Beijing, China) SupermoIII RT kit. qRT-PCR was performed with a DyNAmo Flash SYBR Green qPCR kit (Thermo Fisher Scientific, Waltham, MA, United States), following a previously described protocol (Zhang et al., 2019; Wu et al., 2022). Actin (GenBank number HQ148720) was used as a reference gene (Zhang et al., 2019), and the $2^{-\Delta\Delta CT}$ method was used to determine the gene expression levels.

Conclusion

A comprehensive analysis of the metabolome and transcriptome of two pineapple cultivars during fruit development was performed to explore the dynamic changes in taste-associated metabolites and the molecular mechanisms controlling sweet-acidic taste during pineapple fruit development based on WGCNA. Suc and citrate were the predominant components of sugars and organic acids in mature fruits of both cultivars. At least 15 candidate genes related to sugar metabolism and accumulation were first identified in pineapple, and their expression levels were significantly related to Suc accumulation, including four *INHs* and three *INVs* related to irreversible hydrolysis of Suc, one *HK* and two *PFKs* participating in the Suc–Suc cycle in the cytoplasm, and two *SUTs* and one *SWEET* associated with sugar transporters. As for the metabolism and accumulation of organic acids, 19 candidate genes were first identified in pineapple, and their expression levels were significantly correlated with citric acid accumulation, including four *PKs*, one *PDH*, two *CSs*, one *ACO* and one *NADP-IDH* related to citrate synthesis, two *GADs* related to citrate degradation, two *MDHs* related to malate metabolism, and four H^+-ATPases, one *ALMT* and one *NaDC* associated with acid transport. The key metabolic pathways and gene network for sugar and organic acid accumulation in developing pineapple fruit are shown in **Figures 7**, **8**.

Data availability statement

The original contributions presented in this study are publicly available. This data can be found here: NCBI, PRJNA851558.

Author contributions

XZ and AG supervised the research. YG, YY, and XC prepared the samples and performed the experiments. YG, JW, and QW analyzed the data, wrote, and revised the manuscript. SL provided the materials. All authors contributed to the article and approved the submitted version.

2 http://pineapple.angiosperms.org/pineapple/html/index.html

Funding

This research was supported by National Key R&D Program of China (No. 2018YFD1000504) and Guangdong Provincial Special Fund for Modern Agriculture Industry Technology Innovation Teams (No. 2022KJ109).

Acknowledgments

We acknowledge TopEdit LLC for the linguistic editing and proofreading during the preparation of this manuscript.

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.971506/full#supplementary-material

SUPPLEMENTARY FIGURE 1
OPLS-DA model plots and loading plots for the M1 vs. M2 **(A)**, M1 vs. M3 **(B)**, M1 vs. M4 **(C)**, M2 vs. M3 **(D)**, M2 vs. M4 **(E)**, M3 vs. M4 **(F)**, B1 vs. B2 **(G)**, B1 vs. B3 **(H)**, B2 vs. B3 **(I)** during fruit development of "MD-2" and "Comte de Paris".

References

Ali, M. M., Hashim, N., Aziz, S. A., and Lasekan, O. (2020). Pineapple (*Ananas comosus*): A comprehensive review of nutritional values, volatile compounds, health benefits, and potential food products. *Food Res. Int.* 137:109675. doi: 10.1016/j.foodres.2020.109675

Antony, E., Taybi, T., Courbot, M., Mugford, S. T., Smith, J. A. C., and Borland, A. M. (2008). Cloning, localization and expression analysis of vacuolar sugar transporters in the CAM plant *Ananas comosus* (pineapple). *J. Exp. Bot.* 59, 1895–1908. doi: 10.1093/jxb/ern077

Arampath, P. C., and Dekker, M. (2019). Bulk storage of mango (*Mangifera indica* L.) and pineapple (*Ananas comosus* L.) pulp: Effect of pulping and storage temperature on phytochemicals and antioxidant activity. *J. Sci. Food Agric.* 99, 5157–5167. doi: 10.1002/jsfa.9762

Aslam, M., Deng, L., Wang, X. B., and Wang, Y. (2019). Expression patterns of genes involved in sugar metabolism and accumulation during peach fruit development and ripening. *Sci. Hortic.* 257:108633. doi: 10.1016/j.scienta.2019.108633

Batista-Silva, W., Nascimento, V. L., Medeiros, D. B., Nunes-Nesi, A., Ribeiro, D. M., Zsögön, A., et al. (2018). Modifications in organic acid profiles during fruit development and ripening: Correlation or causation? *Front. Plant Sci.* 9:1689. doi: 10.3389/fpls.2018.01689

Chen, W., Gong, L., Guo, Z. L., Wang, W. S., Zhang, H. Y., Liu, X. Q., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: Application in the study of rice metabolomics. *Mol. Plant* 6, 1769–1780. doi: 10.1093/mp/sst080

Chen, X. D., Cai, W. J., Xia, J., Yu, H. M., Wang, Q. L., Pang, F. H., et al. (2020). Metabolomic and transcriptomic analyses reveal that blue light promotes chlorogenic acid synthesis in strawberry. *J. Agric. Food Chem.* 68, 12485–12492. doi: 10.1021/acs.jafc.0c05020

Dai, Y., Sun, X., Wang, C. G., Li, F., Zhang, S. F., Zang, H., et al. (2021). Gene co-expression network analysis reveals key pathways and hub genes in Chinese cabbage (*Brassica rapa* L.) during vernalization. *BMC Genom.* 22:236. doi: 10.1186/s12864-021-07510-8

Etienne, A., Génard, M., Lobit, P., Mbeguié-A-Mbéguié, D., and Bugaud, C. (2013). What controls fleshy fruit acidity? A review of malate and citrate accumulation in fruit cells. *J. Exp. Bot.* 64, 1451–1469. doi: 10.1093/jxb/ert035

Feng, G. Z., Wu, J. X., Xu, Y. H., Lu, L. Q., and Yi, H. L. (2021). High-spatiotemporal-resolution transcriptomes provide insights into fruit development and ripening in *Citrus sinensis*. *Plant Biotechnol. J.* 19, 1337–1353. doi: 10.1111/pbi.13549

Guo, C. Y., Li, H. Y., Xia, X. Y., Liu, X. Y., and Yang, L. (2018). Functional and evolution characterization of sweet sugar transporters in *Ananas comosus*. *Biochem. Biophys. Res. Commun.* 496, 407–414. doi: 10.1016/j.bbrc.2018.01.024

Hong, K. Q., Chen, L., Gu, H., Zhang, X. M., Chen, J., Nile, S. H., et al. (2021). Novel insight into the relationship between metabolic profile and fatty acid accumulation altering cellular lipid content in pineapple fruits at different stages of maturity. *J. Agric. Food Chem.* 69, 8578–8589. doi: 10.1021/acs.jafc.1c02658

Huang, X. Y., Wang, C. K., Zhao, Y. W., Sun, C. H., and Hu, D. G. (2021). Mechanisms and regulation of organic acid accumulation in plant vacuoles. *Hortic. Res.* 8:227. doi: 10.1038/s41438-021-00702-z

Igamberdiev, A. U., and Eprintsev, A. T. (2016). Organic acids: The pools of fixed carbon involved in redox regulation and energy balance in higher plants. *Front. Plant Sci.* 7:1042. doi: 10.3389/fpls.2016.01042

Ikram, M. M. M., Ridwani, S., Putri, S. P., and Fukusaki, E. (2020). GC-MS based metabolite profiling to monitor ripening-specific metabolites in pineapple (*Ananas comosus*). *Metabolites* 10:134. doi: 10.3390/metabo10040134

Kawaguchi, K., Takei-Hoshi, R., Yoshikawa, I., Nishida, K., Kobayashi, M., Kusano, M., et al. (2021). Functional disruption of cell wall invertase inhibitor by genome editing increases sugar content of tomato fruit without decrease fruit weight. *Sci. Rep.* 11:21534. doi: 10.1038/s41598-021-00966-4

Klee, H. J., and Tieman, D. M. (2018). The genetics of fruit flavour preferences. *Nat. Rev. Genet.* 19, 347–356. doi: 10.1038/s41576-018-0002-5

Langfelder, P., and Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559

Léchaudel, M., Darnaudery, M., Joët, T., Fournier, P., and Joas, J. (2018). Genotypic and environmental effects on the level of ascorbic acid, phenolic compounds and related gene expression during pineapple fruit development and ripening. *Plant Physiol. Biochem.* 130, 127–138. doi: 10.1016/j.plaphy.2018.06.041

Li, J. M., Qin, M. F., Qiao, X., Cheng, Y. S., Li, X. L., Zhang, H. P., et al. (2017). A new insight into the evolution and functional divergence of SWEET transporters in Chinese white pear (*Pyrus bretschneideri*). *Plant Cell Physiol.* 58, 839–850. doi: 10.1093/pcp/pcx025

Li, M. J., Feng, F. J., and Cheng, L. L. (2012). Expression patterns of genes involved in sugar metabolism and accumulation during apple fruit development. *PLoS One* 7:e33055. doi: 10.1371/journal.pone.0033055

Liao, G. L., Liu, Q., Xu, X. B., He, Y. Q., Li, Y. Q., Wang, H. L., et al. (2021). Metabolome and transcriptome reveal novel formation mechanism of early mature trait in kiwifruit (*Actinidia eriantha*). *Front. Plant Sci.* 12:760496. doi: 10.3389/fpls.2021.760496

Liu, X., Hu, X. M., Jin, L. F., Shi, C. Y., Liu, Y. Z., and Peng, S. A. (2014). Identification and transcript analysis of two glutamate decarboxylase genes, CsGAD1 and CsGAD2, reveal the strong relationship between CsGAD1 and citrate utilization in citrus fruit. *Mol. Biol. Rep.* 41, 6253–6262. doi: 10.1007/s11033-014-3506-x

Lu, X. H., Sun, D. Q., Wu, Q. S., Liu, S. H., and Sun, G. M. (2014). Physico-chemical properties, antioxidant activity and mineral contents of pineapple genotypes grown in China. *Molecules* 19, 8518–8532. doi: 10.3390/molecules19068518

Ma, M., Wang, L. B., Zhang, S. L., Guo, L., Zhang, Z., Li, J., et al. (2020). Acid vacuolar invertase 1 (PbrAc-Inv1) and invertase inhibitor 5 (PbrII5) were involved in sucrose hydrolysis during postharvest pear storage. *Food Chem.* 320:126635. doi: 10.1016/j.foodchem.2020.126635

McCommis, K. S., and Finck, B. N. (2015). Mitochondrial pyruvate transport: A historical perspective and future research directions. *Biochem. J.* 466, 443–454. doi: 10.1042/BJ20141171

Minami, A., Kang, X. J., and Carter, C. J. (2021). A cell wall invertase controls nectar volume and sugar composition. *Plant J.* 107, 1016–1028. doi: 10.1111/tpj.15357

Ming, R., VanBuren, R., Wai, C. M., Tang, H. B., Schatz, M. C., Bowers, J. E., et al. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nat. Genet.* 47, 1435–1442. doi: 10.1038/ng.3435

Miron, D., Petreikov, M., Carmi, N., Shen, S., Levin, I., Granot, D., et al. (2002). Sucrose Uptake, invertase localization and gene expression in developing fruit of *Lycopersicon esculentum* and the sucrose-accumulating *Lycopersicon hirsutum*. *Physiol. Plant.* 115, 35–47. doi: 10.1034/j.1399-3054.2002.1150104.x

Peng, Q., Cai, Y. M., Lai, E. H., Nakamura, M., Liao, L., Zheng, B. B., et al. (2020). The sucrose transporter MdSUT4.1 participates in the regulation of fruit sugar accumulation in apple. *BMC Plant Biol.* 20:191. doi: 10.1186/s12870-020-02406-3

Ru, L., He, Y., Zhu, Z. J., Patrick, J. W., and Ruan, Y. L. (2020). Integrating sugar metabolism with transport: Elevation of endogenous cell wall invertase activity up-regulates SlHT2 and SlSWEET12c expression for early fruit development in tomato. *Front. Genet.* 11:592596. doi: 10.3389/fgene.2020.592596

Ruan, Y. L., Jin, Y., Yang, Y. J., Li, G. J., and Boyer, J. S. (2010). Sugar input, metabolism, and signaling mediated by invertase: Roles in development, yield potential, and response to drought and heat. *Mol. Plant* 3, 942–955. doi: 10.1093/mp/ssq044

Saradhuldhat, P., and Paull, R. E. (2007). Pineapple organic acid metabolism and accumulation during fruit development. *Sci. Hortic.* 112, 297–303. doi: 10.1016/j.scienta.2006.12.031

Shi, C. Y., Hussain, S. B., Yang, H., Bai, Y. X., Khan, M. A., and Liu, Y. Z. (2019). CsPH8, a P-type proton pump gene, plays a key role in the diversity of citric acid accumulation in citrus fruits. *Plant Sci.* 289:110288. doi: 10.1016/j.plantsci.2019.110288

Shi, C. Y., Song, R. Q., Hu, X. M., Liu, X., Jin, L. F., and Liu, Y. Z. (2015). Citrus PH5-like H+-ATPase genes: Identification and transcript analysis to investigate their possible relationship with citrate accumulation in fruits. *Front. Plant Sci.* 6:135. doi: 10.3389/fpls.2015.00135

Steingass, C. B., Carle, R., and Schmarr, H. G. (2015). Ripening-dependent metabolic changes in the volatiles of pineapple (*Ananas comosus* (L.) Merr.) fruit: I. Characterization of pineapple aroma compounds by comprehensive two-dimensional gas chromatography-mass spectrometry. *Anal. Bioanal. Chem.* 407, 2591–2608. doi: 10.1007/s00216-015-8475-y

Strazzer, P., Spelt, C. E., Li, S. J., Bliek, M., Federici, C. T., Roose, M. L., et al. (2019). Hyperacidification of Citrus fruits by a vacuolar proton-pumping P-ATPase complex. *Nat. Commun.* 10:744. doi: 10.1038/s41467-019-08516-3

Umer, M. J., Safdar, L. B., Gebremeskel, H., Zhao, S. J., Yuan, P. L., Zhu, H. J., et al. (2020). Identification of key gene networks controlling organic acid and sugar metabolism during watermelon fruit development by integrating metabolic phenotypes and gene expression profiles. *Hortic. Res.* 7:193. doi: 10.1038/s41438-020-00416-8

Vimolmangkang, S., Zheng, H. Y., Peng, Q., Jiang, Q., Wang, H. L., Fang, T., et al. (2016). Assessment of sugar components and genes involved in the regulation of sucrose accumulation in peach fruit. *J. Agric. Food Chem.* 64, 6723–6729. doi: 10.1021/acs.jafc.6b02159

Walker, R. P., Bonghi, C., Varotto, S., Battistelli, A., Burbidge, C. A., Castellarin, S. D., et al. (2021). Sucrose metabolism and transport in grapevines, with emphasis on berries and leaves, and insights gained from a cross-species comparison. *Int. J. Mol. Sci.* 22:7794. doi: 10.3390/ijms22157794

Wang, L. L., Li, Y., Jin, X. Y., Liu, L. P., Dai, X. Z., Liu, Y. H., et al. (2020). Floral transcriptomes reveal gene networks in pineapple floral growth and fruit development. *Commun. Biol.* 3:500. doi: 10.1038/s42003-020-01235-2

Wang, R., Shu, P., Zhang, C., Zhang, J. L., Chen, Y., Zhang, Y. X., et al. (2022). Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytol.* 233, 373–389. doi: 10.1111/nph.17618

Wang, X., Chen, Y., Jiang, S., Xu, F., Wang, H. F., Wei, Y. Y., et al. (2020). PpINH1, an invertase inhibitor, interacts with vacuolar invertase PpVIN2 in regulating the chilling tolerance of peach fruit. *Hortic. Res.* 7:168. doi: 10.1038/s41438-020-00389-8

Wen, S. Y., Neuhaus, H. E., Cheng, J. T., and Bie, Z. L. (2022). Contributions of sugar transporters to crop yield and fruit quality. *J. Exp. Bot.* 73, 2275–2289. doi: 10.1093/jxb/erac043

Wu, J. Y., Chen, M., Yao, Y. L., Fu, Q., Zhu, Z. Y., and Zhang, X. M. (2022). Identification, characterisation, and expression profile analysis of the sucrose phosphate synthase gene family in pineapple (*Ananas comosus*). *J. Hortic. Sci. Biotechnol.* 97, 201–210. doi: 10.1080/14620316.2021.1981778

Xu, H. M., Yu, Q. Y., Shi, Y., Hua, X. T., Tang, H. B., Yang, L., et al. (2018). PGD: Pineapple genomics database. *Hortic. Res.* 5:66. doi: 10.1038/s41438-018-0078-2

Yuan, H. Z., Pang, F. H., Cai, W. J., Chen, X. D., Zhao, M. Z., and Yu, H. M. (2021). Genome-wide analysis of the invertase genes in strawberry (*Fragaria* × *ananassa*). *J. Integr. Agric.* 20, 2652–2665. doi: 10.1016/S2095-3119(20)63381-0

Zanon, L., Falchi, R., Hackel, A., Kühn, C., and Vizzotto, G. (2015a). Expression of peach sucrose transporters in heterologous systems points out their different physiological role. *Plant Sci.* 238, 262–272. doi: 10.1016/j.plantsci.2015.06.014

Zanon, L., Falchi, R., Santi, S., and Vizzotto, G. (2015b). Sucrose transport and phloem unloading in peach fruit: Potential role of two transporters localized in different cell types. *Physiol. Plant.* 154, 179–193. doi: 10.1111/ppl.12304

Zhang, Q., Liu, Y. L., He, C. C., and Zhu, S. J. (2015). Postharvest exogenous application of abscisic acid reduces internal browning in pineapple. *J. Agric. Food Chem.* 63, 5313–5320. doi: 10.1021/jf506279x

Zhang, X. M., Liu, S. H., Du, L. Q., Yao, Y. L., and Wu, J. Y. (2019). Activities, transcript levels, and subcellular localizations of sucrose phosphate synthase, sucrose synthase, and neutral invertase and change in sucrose content during fruit development in pineapple (*Ananas comosus*). *J. Hortic. Sci. Biotechnol.* 94, 573–579. doi: 10.1080/14620316.2019.1604169

Zhang, X. M., Wang, W., Du, L. Q., Xie, J. H., Yao, Y. L., and Sun, G. M. (2012). Expression patterns, activities and carbohydrate-metabolizing regulation of sucrose phosphate synthase, sucrose synthase and neutral invertase in pineapple fruit during development and ripening. *Int. J. Mol. Sci.* 13, 9460–9477. doi: 10.3390/ijms13089460

Zhang, X. S., Feng, C. Y., Wang, M. N., Li, T. L., Liu, X., and Jiang, J. (2021). Plasma membrane-localized SlSWEET7a and SlSWEET14 regulate sugar transport and storage in tomato fruits. *Hortic. Res.* 8:186. doi: 10.1038/s41438-021-00624-w

Zheng, B. B., Zhao, L., Jiang, X. H., Cherono, S., Liu, J. J., Ogutu, C., et al. (2020). Assessment of organic acid accumulation and its related genes in peach. *Food Chem.* 334:127567. doi: 10.1016/j.foodchem.2020.127567

Zhu, L. C., Li, B. Y., Wu, L. M., Li, H. X., Wang, Z. Y., Wei, X. Y., et al. (2021). MdERDL6-mediated glucose efflux to the cytosol promotes sugar accumulation in the vacuole through up-regulating TSTs in apple and tomato. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2022788118. doi: 10.1073/pnas.2022788118

Zou, S. C., Wu, J. C., Shahid, M. Q., He, Y. H., Lin, S. Q., Liu, Z. H., et al. (2020). Identification of key taste components in loquat using widely targeted metabolomics. *Food Chem.* 323:126822. doi: 10.1016/j.foodchem.2020.126822

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 11 October 2022
DOI 10.3389/fpls.2022.943547

OPEN ACCESS

EDITED BY
Shunquan Lin,
South China Agricultural University, China

REVIEWED BY
Dong Wang,
Agricultural Research Service (USDA), United States
Hongcheng Fang,
Shandong Agricultural University, China

*CORRESPONDENCE
Zhaohe Yuan
zhyuan88@hotmail.com

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 13 May 2022
ACCEPTED 20 September 2022
PUBLISHED 11 October 2022

CITATION
Wang Y, Zhao Y, Wu Y, Zhao X, Hao Z, Luo H and Yuan Z (2022) Transcriptional profiling of long non-coding RNAs regulates fruit cracking in *Punica granatum* L. under bagging. *Front. Plant Sci.* 13:943547. doi: 10.3389/fpls.2022.943547

Transcriptional profiling of long non-coding RNAs regulating fruit cracking in *Punica granatum* L. under bagging

Yuying Wang[1], Yujie Zhao[1], Yaqiong Wu[2], Xueqing Zhao[1], Zhaoxiang Hao[3], Hua Luo[3] and Zhaohe Yuan[1]*

[1]Co-Innovation Center for Sustainable Forestry in Southern China, College of Forestry, Nanjing Forestry University, Nanjing, China, [2]Institute of Botany, Jiangsu Province and Chinese Academy of Sciences, Nanjing, China, [3]Zaozhuang Pomegranate Research Center, Institute of Botany, Zaozhuang, China

Fruit cracking tremendously damages the appearance of fruit, easily leads to pathogen invasion, greatly reduces the marketability and causes immense economic losses. The pivotal role of long non-coding RNAs (lncRNAs) in diverse biological processes has been confirmed, while the roles of lncRNAs underlying fruit cracking remain poorly understood. In this study, the incidence of fruit cracking was 7.26% under the bagging treatment, the control group was 38.11%, indicating that bagging considerably diminished the fruit cracking rate. LncRNA libraries for fruit cracking (FC), fruit non-cracking (FNC) and fruit non-cracking under bagging (FB) in pomegranate (*Punica granatum* L.) were performed and analysed *via* high-throughput transcriptome sequencing. A total of 3194 lncRNAs were obtained with a total length of 4898846 nt and an average length of 1533.77 nt in pomegranate. We identified 42 differentially expressed lncRNAs (DELs) and 137 differentially expressed mRNAs (DEGs) in FC vs FNC and 35 DELs and 160 DEGs in FB vs FC that formed co-expression networks respectively, suggesting that there are involved in phytohormone signaling pathway, lignin catabolic process, lipid transport/binding, cutin biosynthetic process and cell wall organization. We also found that 18 *cis*-acting DELs regulated 18 target genes, and 10 *trans*-acting DELs regulated 24 target genes in FC vs FNC, 23 DELs regulate 23 target genes for the cis-acting lncRNAs and 12 DELs regulated 36 target genes in FB vs FC, which provides an understanding for the regulation of the fruit cracking. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis results demonstrated that DELs participated in calcium ion binding, glycerophospholipid metabolism, flavonoid biosynthetic process, cell wall biogenesis, xyloglucan metabolic process, hormone signal transduction and starch and sucrose metabolism. Our findings provide new insights into the roles of lncRNAs in regulating the fruit cracking and lay the foundation for further improvement of pomegranate quality.

KEYWORDS

pomegranate, bagging, lncRNA, target genes, functional analysis, fruit cracking

Introduction

Pomegranate (*Punica granatum* L.) is a deciduous shrub or small tree of genus *Punica* in the Lythraceae family (Yuan et al., 2017; Lyu et al., 2020) and it is native to Central Asia such as Iran, Afghanistan and the Caucasus (Qin et al., 2017). Pomegranate possesses an outstanding ornamental and medicinal value and has become an emerging functional fruit due to the richness of ellagitannin-based compounds in the peel and aril, which have antioxidant, antibacterial, anti-inflammatory, anticancer, antidiabetic and cardiovascular health promoting properties (Vlachojannis et al., 2015; Baghel et al., 2021). At present, the research on pomegranate mainly focuses on evolution, ovule, flower, seed coat, fruit development (Saminathan et al., 2016; Yuan et al., 2017; Zhao et al., 2020; Qin et al., 2020), natural product biosynthesis (Ono et al., 2011), pericarp color synthesis (Zhao et al., 2015) and salt stress response (Liu et al., 2020). Few studies have focused on the quality of pomegranate peel appearance and its regulation mechanism.

Fruit cracking affects its appearance and quality, resulting in a decrease in the commercial value and marketability, resulting in economic loss (Ginzberg and Stern, 2019). Fruit cracking is an intricate biological process involving phytohormone signaling pathway, lipid metabolism, cutin biosynthetic process and cell wall organization, etc (Chen et al., 2019; Wang et al., 2019; Jiang et al., 2019). The outer layer of pericarp is covered with cutin and wax, which plays a vital role in resisting pathogens, reducing water loss and enhancing the storage and quality of fruit (Lewandowska et al., 2020; Li et al., 2021; Wang et al., 2022). Previous studies commonly reported to be prone to cracking include pomegranate (Galindo et al., 2014), cherry (Quero-García et al., 2021), grape (Zhang et al., 2021), apple (Joshi et al., 2018), jujube (Li et al., 2020b), Akebia trifoliate (Niu et al., 2020) and tomato (Capel et al., 2017). Fruit cracking is influenced by genetic factors (fruit genotype, cuticle and epidermal cells, cell wall, etc.) (Cortés et al., 1983; Correia et al., 2020), environmental factors (humidity, light and temperature, etc.) (Wang and Camp, 2000; Galindo et al., 2014), cultivation management practices (irrigating, bagging and spraying of plant growth regulators and minerals, etc.) (Joshi et al., 2018; Xue et al., 2020; Li et al., 2020c; Correia et al., 2020) and postharvest storage factors (hydrocooling and postharvest calcium treatment, etc.) (Wang and Long, 2015; Singh et al., 2021). The pomegranate fruit cracking was closely related to fruit volume and shape (Sharifani, 2014). Galindo et al. (Galindo et al., 2014) suggested that pomegranate fruits were sensitive to water deficit at the end of fruit growth and ripening period. Rainfall can greatly affect previously water-stressed pomegranate plants, and an increase in aril turgor stresses the pericarp and makes it susceptible to cracking. Joshi et al. (2021) showed that high growth rates weakened the pericarp and increased the susceptibility of the fruit to cracking. Multiple previous studies have shown that bagging immensely reduces the incidence of sunburn and fruit cracking (Griñán et al., 2019; Asrey et al., 2020). Hosein-Beigi et al. (2019) investigated the effect of calcium (Ca) (0, 0.75% and 1.5%), boron (B) (0 and 3000 ppm) and gibberellin (GA_3) (0, 75 and 150 ppm) on 'Malase-Torshe-Saveh' pomegranate fruit, the application of boron can increase the water content of the peel, thereby combination spraying of Ca, B and GA_3 was more effective than spraying alone, with the least symptoms of fruit cracking and sunburn.

Non-coding RNAs (ncRNAs) refers to the general term for RNAs that do not encode proteins, including ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), microRNAs (miRNAs), long non-coding RNAs (lncRNAs), circular RNAs (circRNAs), etc (Kim and Sung, 2012). LncRNAs transcripts are longer than 200 nucleotides and no protein-encoding function (Ulitsky and Bartel, 2013; Laurent et al., 2015). LncRNAs were originally considered to be the 'noise' of genome transcription, without biological function (Ulitsky and Bartel, 2013; Fatica and Bozzoni, 2014; Chekanova, 2015). Recent studies have shown that lncRNAs can regulate gene expression at epigenetic, transcriptional and post-transcriptional levels, and are involved in a variety of crucial regulatory processes such as photomorphogenesis, fruit softing and ripening, which are closely related to the growth and development of plants (Hadjiargyrou and Delihas, 2013; Wang et al., 2014; Berry and Dean, 2015; Correa et al., 2018). With the speedy development of high-throughput sequencing, considerable number of lncRNAs have been widely identified in Arabidopsis thaliana (Liu et al., 2012; Zhu et al., 2014), Oryza sativa (Li et al., 2007), Solanum lycopersicum (Zhu et al., 2015; Xue et al., 2020), Medicago truncatula (Wang et al., 2015) and Ginkgo biloba (Wu et al., 2019b). However, the function of lncRNAs in fruit cracking is rarely limited. Thus, it is considerable and urgent to reveal and investigate the function of lncRNAs in fruit cracking of pomegranate.

The pomegranate cultivar used was *Punica granatum* L. cv. 'Daqingpitian', an excellent variety and main planted variety in Zaozhuang City, Shandong Province, accounting for about 80% of the total cultivation. The fruit of 'Daqingpitian' is large-sized, with bright pericarp and wonderful appearance, nevertheless, it is cracking-susceptible. Base on this, we performed high-throughput transcriptome sequencing to reveal lncRNAs involved in pomegranate fruit cracking, and carried out coexpression analysis of differentially expressed lncRNAs (DELs) and differentially expressed mRNAs (DEGs). Furthermore, the cis- and trans-target genes of DELs were explored to expound the putative biological functions. Those results provide a foundation for future research of the molecular mechanism of fruit cracking.

Materials and methods

Plant materials and bagging treatment

Pomegranate cultivars 'Daqingpitian' was used in our experiment. Pomegranate trees that were about 15 years old were selected from an orchard in Zaozhuang City, Shandong Province, China (34°77' N, 117°56' E). Fruits on ten trees were covered with white single-layer bags on 110 days after full bloom (DAFB), 2020, fruits on ten trees without bagging treatment. In addition, these trees were subject to the same orchard management strategies. Types and incidence of fruit cracking were investigated at the final fruit ripening stage. Total fruit numbers for bagging and control are 358 and 307, respectively. We collected samples of fruit from non-cracking (FNC) to fruit cracking (FC) and non-cracking under bagging (FB) during late ripening. Three non-cracking fruits from three pomegranate trees with the same direction of light were selected as a biological replicate, the cracking fruits, non-cracking fruits under bagging were the same as that of non-cracking fruits and three biological replicates of each group of samples. The pericarps were immediately isolated into 15 ml centrifuge tubes, snap frozen in liquid nitrogen for 10 min and stored at -80°C.

RNA extraction and transcriptome sequencing

The total RNA from these pericarps (FNC, FC and FB) was extracted using a mirVana miRNA Isolation Kit (Ambion-1561) in accordance with the manufacturer's instructions. The RNA integrity, and quality were checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Following analysis was performed on samples with RNA integrity number (RIN)≥ 7. Then, 6 libraries were constructed and sequenced on an Illumina HiSeq^{TM} X platform used for reference transcriptome sequencing. Trimmomatic software (Bolger et al., 2014) was used to filter out adapter, N and low-quality reads contained in raw reads to obtain clean reads. After comparing clean reads to the reference genome (ASM765513v2) using Hisat2 (Kim et al., 2019) and obtaining a bam file of the comparing results, the reads comparing to the genes were subsequently assembled using StringTie (Pertea et al., 2015) and the individually assembled transcripts from each sample were merged a complete transcript. The raw data was deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra/) under accession number PRJNA773212.

Identification of lncRNAs

The lncRNAs can be classified into intergenic lncRNA (lincRNA), intronic lncRNA, anti-sense lncRNA, sense-lncRNA, bidirectional-lncRNA and other types according to the position relationship between lncRNA and coding sequence. Based on the characteristics of lncRNAs, the candidate lncRNAs were obtained by a rigorous four-step screening method: (1) The cuffcompare software (Ghosh and Chan, 2016) was used to compare merged transcripts with reference transcripts one by one to identify the location types of the remaining transcripts. Then, through the screening of candidate lncRNA transcripts, the transcripts with 'I', 'u', 'x', and 'o' were reserved. (2) Transcripts were screened according to the length of more than 200 nt and the number of exons greater than or equal to 2. (3) The transcripts obtained were analyzed for coding ability prediction using four software, CPC2 (Kang et al., 2017), CNCI (Sun et al., 2013), Pfam (Sonnhammer et al., 1998) and PLEK (Li et al., 2014), and transcripts with coding potential were screened out. (4) For the species with known lncRNA, the lncRNA sequences predicted in the third step were aligned with the known lncRNA sequences by using BLASTN software. Quantitative analysis was performed after combining with known lncRNA sequences. For species with no known lncRNA, the lncRNA sequences obtained in the third step were directly used for quantitative analysis.

Expression level quantification and DEG, DEL screening

The FPKM method (Roberts et al., 2011) eliminates the effect of transcript length and sequencing amount differences on the calculation of transcript expression, this method was applied to calculate the transcript expression level. Then, DESeq software (Anders and Huber, 2012) was used to standardize the number of counts of each sample mRNA, lncRNA and calculated the difference multiple, and negative binomial distribution test (NB) was used to test the difference significance of the reads number, and finally the difference mRNAs, lncRNAs were screened based on the difference multiple (|log2 FC| >1) and difference significance test results (P-value< 0.05).

DELs and DEGs localization and coexpression network construction

A plot was generated to display the localization and abundance of DELs and DEGs in the pomegranate genome by Circos software (Krzywinski et al., 2009). Expression correlation was calculated based on DELs and DEGs expression data using Pearson correlation test, and pairs with correlation coefficients of not less than 0.8 and p values less than or equal to 0.05 were selected and considered that there was a co-expression relationship. Cytoscape (Shannon et al., 2003) was used to construct a co-expression network.

Target gene prediction and enrichment analysis

Potential target genes of lncRNAs were predicted according to different regulatory patterns in cis or trans-acting, respectively. The principle of cis-target gene prediction is that the function of lncRNAs is related to their neighboring protein-coding genes. LncRNAs located upstream may intersect with promoters or other cis-acting elements of co-expressed genes, thus regulating gene expression at the transcriptional or post-transcriptional level; lncRNAs located in the 3'UTR or downstream of genes may be involved in other regulatory roles. All coding genes that were near the lncRNAs in the up- or down-stream 100 kb and significantly co-expressed with the lncRNA were defined as target genes. The function of trans-acting target genes for lncRNAs is not related to the location of the coding genes, but rather to the protein-coding genes they co-express. The target genes were predicted by the correlation analysis method between the expression of lncRNAs and protein-coding genes. To preferably comprehend functions of the lncRNA target genes, GO and KEGG enrichment analyses were carried out.

Quantitative real-time PCR (qRT-PCR) validation

Ten DELs were randomly selected from the outcomes of the transcriptional research and qRT-PCR was performed on the Applied Biosystems 7500. Specific primers of ten DELs were designed (Table S1). The Actin of pomegranate (F: AGTCCTCTTCCAGCCATCTC and R: CACTGAGCACA ATGTTTCCA) was used as the internal reference gene. All reactions were carried out in three biological replicates. The Amplification procedure of qRT-PCR was 95°C for 5 min; 95°C for 15 s, 58°C for 45 s, a total of 40 cycles. The relative expression level was normalized by $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Results

Types and incidence of fruit cracking

The four types of fruit cracking in pomegranate was investigated, including longitudinal cracking, transversal cracking, longitudinal and transversal cracking and irregular cracking (Figure 1). Among them, the longitudinal cracking ratio was the highest about 70% (The number of longitudinal cracking fruits and total are 19 and 26 in fruit bagging, and 81 and 117 in control), and the other three were about 30%. We calculated that the incidence of fruit cracking under the bagging treatment was 7.26%, while the control group was as high as 38.11% (Table 1), indicating that bagging substantially reduced the fruit cracking rate.

RNA sequencing and identification of lncRNAs

In total, we obtained a total of 287.43 M, 282.83 M and 280.02 M raw reads from pomegranate fruit non-cracking under bagging (FB), fruit cracking (FC), fruit non-cracking (FNC), respectively (Table 2). After filtering out low quality sequences, approximately 284.96M, 280.5 M and 277.65 clean reads were generated from FB, FC and FNC, Q30 bases were distributed in 95.04~95.57%, with an average GC content of 50.08%. In addition, through coding potential calculator (CPC), coding-non-coding index (CNCI), protein families (Pfam) and predictor of long noncoding RNAs and messenger RNAs based on an improved k-mer scheme (PLEK) analysis, 3194 lncRNAs were identified with a total length of 4898846 nt and an average length of 1533.77 nt. The length distribution of lncRNAs showed that the maximum number of lncRNAs greater than 2000 bp was 701, accounting for 22% of the total number, and the minimum number of lncRNAs less than or equal to 200 bp was only 2 (Figure 2A). In addition, through the positional relationship between lncRNAs

FIGURE 1
Types of fruit cracking.

10.3389/fpls.2022.943547

TABLE 1 The incidence of fruit cracking under the bagging treatment and control.

Treatment	No.	Total	Number of cracked fruits	Fruit cracking rates
Fruit bagging	1	52	6	11.54%
	2	41	5	12.20%
	3	28	3	10.71%
	4	49	2	4.08%
	5	31	0	0
	6	36	0	0
	7	33	7	21.21%
	8	34	2	5.88%
	9	26	1	3.85%
	10	28	0	0
		358	26	7.26%
Control	1	51	28	54.90%
	2	18	5	27.78%
	3	21	10	47.62%
	4	14	1	7.14%
	5	24	10	41.67%
	6	21	2	9.52%
	7	25	11	44.00%
	8	29	18	62.07%
	9	48	20	41.67%
	10	56	12	21.43%
		307	117	38.11%

and known protein-coding transcripts, lncRNA types are counted from three levels, namely direction (sense and antisense), type (intergenic and genic), location (upstream, downstream, exonic and intronic). Among them, the direction was sense, and the number of lncRNAs in intergenic-upstream was at most 616 (Figure 2B). The exons number of lncRNAs indicated that the majority of lncRNAs contained 2 exons (Figure 2C).

LncRNAs expression level analysis

The fragments per kilobase per million reads (FPKM) was used to evaluate expression levels of 3194 lncRNAs transcripts in pomegranate FB, FC and FNC (Table S2). The overall expression levels of lncRNAs were different between FB, FC and FNC (Figure 3A). In addition, the results indicated differences in transcript numbers and distributions of gene expression values (Figure 3B). Approximately FPKM values of half of the transcripts were between 0 and 0.5 and about FPKM values of one-third of the transcripts were greater than or equal to 10.

Differential expression and localization analysis

A total of 61 DELs were identified with a P-value< 0.05 and | log2 FC| > 1 in FC vs FNC (Table S3), 62 DELs in FB vs FC (Table S4). Among them, 31 were up-regulated and 30 were down-regulated in the FC vs FNC, 18 were up-regulated and 44 were down-regulated in the FB vs FC, as shown in a volcano plot (Figures 4A, B). Also, the expression patterns of the FC vs FNC, FB vs FC were compared, red and blue revealed the highest and lowest expression level, respectively. The heat map clustering showed that FC vs FNC, FB vs FC appear in the same cluster by hierarchical clustering analysis, indicating the genes in the same cluster may have similar biological functions (Figures 4C, D).

To display the position of DELs and DEGs more intuitively, Circos software was used to show them (Figure 5). The Circos diagram clearly revealed that the distribution of DELs and DEGs on chromosomes was not uniform. Among FC vs FNC, DELs were less distributed on chromosome 3 and chromosome 7, and were the least distributed on chromosome 6 (2), the same as FB vs FC; DELs were most distributed on chromosome 6 (12) (Figure 5A). In addition, the number of DELs distributed on chromosome 4 were the largest 12 in FB vs FC, followed by chromosome 5 (Figure 5B).

TABLE 2 Quality of the sequencing data.

Sample	Raw Reads	Raw Bases	Clean Reads	Clean Bases	Valid Bases	Q30	GC
FB1	95.56M	14.33G	94.74M	13.01G	90.77%	95.41%	49.90%
FB2	93.34M	14.00G	92.53M	12.79G	91.37%	95.12%	49.64%
FB3	98.53M	14.78G	97.69M	13.55G	91.68%	95.04%	50.45%
FC1	95.65 M	14.35 G	94.85 M	12.98 G	90.44%	95.44%	50.07%
FC2	94.00 M	14.10 G	93.25 M	12.95 G	91.82%	95.44%	50.18%
FC3	93.18 M	13.98 G	92.40 M	12.79 G	91.53%	95.36%	50.22%
FNC1	96.12 M	14.42 G	95.32 M	13.15 G	91.18%	95.37%	50.28%
FNC2	92.18 M	13.83 G	91.42 M	12.70 G	91.82%	95.57%	50.12%
FNC3	91.72 M	13.76 G	90.91 M	12.65 G	91.93%	95.20%	49.85%

10.3389/fpls.2022.943547

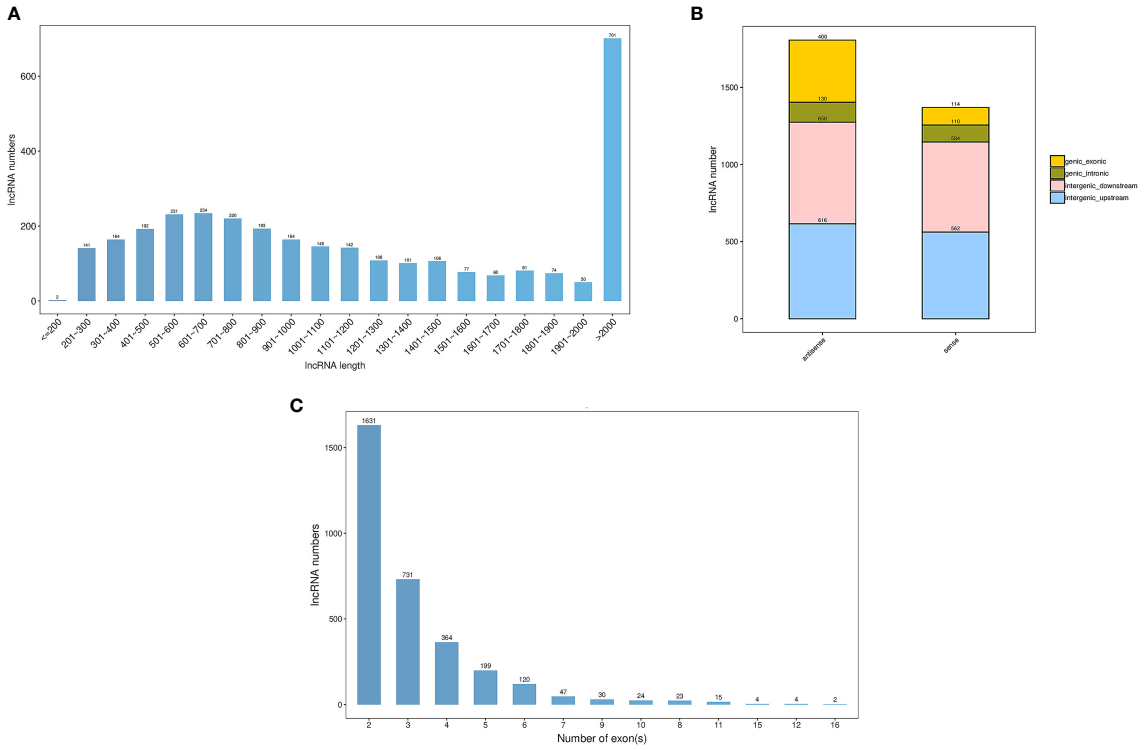


FIGURE 2
Summary of the characteristic long noncoding RNAs (lncRNAs) in pomegranate. **(A)** The length distribution of lncRNAs, **(B)** The classification of lncRNAs and **(C)** The exon number of lncRNAs.

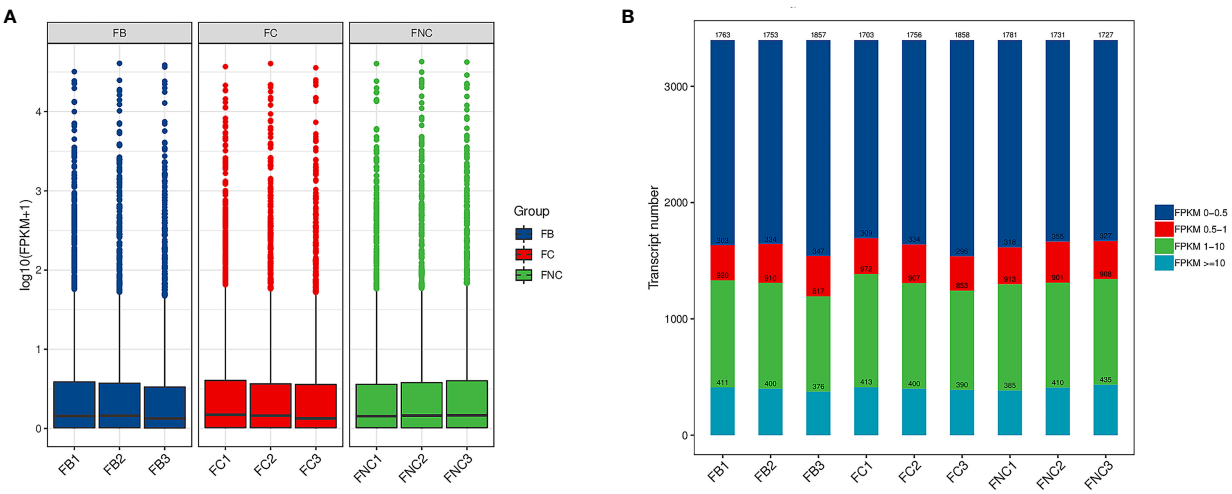


FIGURE 3
Long noncoding RNAs (lncRNAs) expression level analysis in pomegranate. **(A)** Boxplot of fragments per kilobase per million reads (FPKM) values. **(B)** Transcript expression b.

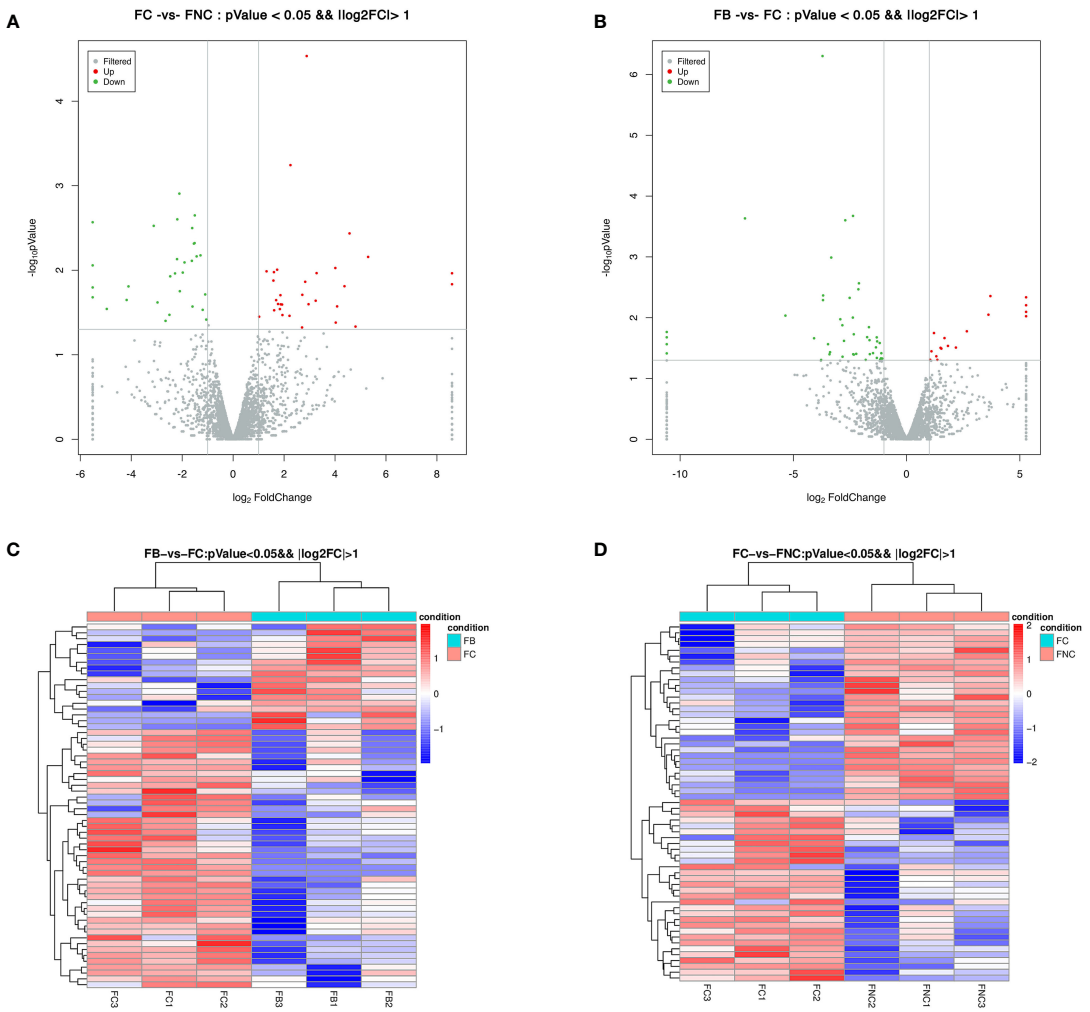


FIGURE 4
Differentially expressed lncRNAs (DELs) in pomegranate. Volcano plot of DELs in FC vs FNC **(A)** and FB vs FC **(B)**. Red and green dots represent significantly up-regulated and significantly down-regulated genes, respectively. Heat map of DELs based on hierarchical clustering analysis in FC vs FNC **(C)** and FB vs FC **(D)**. Red and blue indicate the highest and lowest expression level of DELs.

Coexpression analysis of DELs and DEGs

To further explore the temporal and spatial transcription patterns and biological functions between DELs and DEGs, a coexpression network was constructed. The results suggested that a total of 42 DELs and 137 DEGs formed a coexpression network in FC vs FNC (Figure 6). Network analysis showed that one DEL co-regulated with one DEG or several DEGs and vice versa. It is suggested that there is a complex interaction between DELs and DEGs. Among them, XR_004158164.1 co-regulated with ten DEGs (LOC116201766, LOC116202352, LOC116198068, LOC116202713, LOC116213464, LOC116212471, LOC116212884, LOC116206401, LOC116211848 and LOC116200280) involved in ethylene-activated signaling pathway, lignin catabolic process, lipid transport/binding, cutin biosynthetic process and cell wall. XR_004156395.1 co-regulated with six DEGs (LOC116214125, LOC116208411, LOC116212683, LOC116210477, LOC116207007 and LOC116213296) involved in lipid transport/binding, cutin biosynthetic process and cell wall. TCONS_00001938 co-regulated with LOC116197025 (Polygalacturonase-1, PG1) involved in lipid transport, wax biosynthetic process and cell wall organization. In addition, XR_004157290.1 co-regulated with LOC116206904 (E3 ubiquitin-protein ligase-23, PUB23) and LOC116197199 (Polygalacturonase-2, PG2) involved in phytohormone signaling pathways, including ethylene-activated/abscisic acid- activated/jasmonic acid mediated signaling pathway.

Among FB vs FC, 35 DELs and 160 DEGs constituted coexpression network (Figure 7). Among them, TCONS_00009904 co-regulated with 17 DEGs involved in lignin catabolic/biosynthetic process, secondary metabolic process, suberin biosynthetic process,

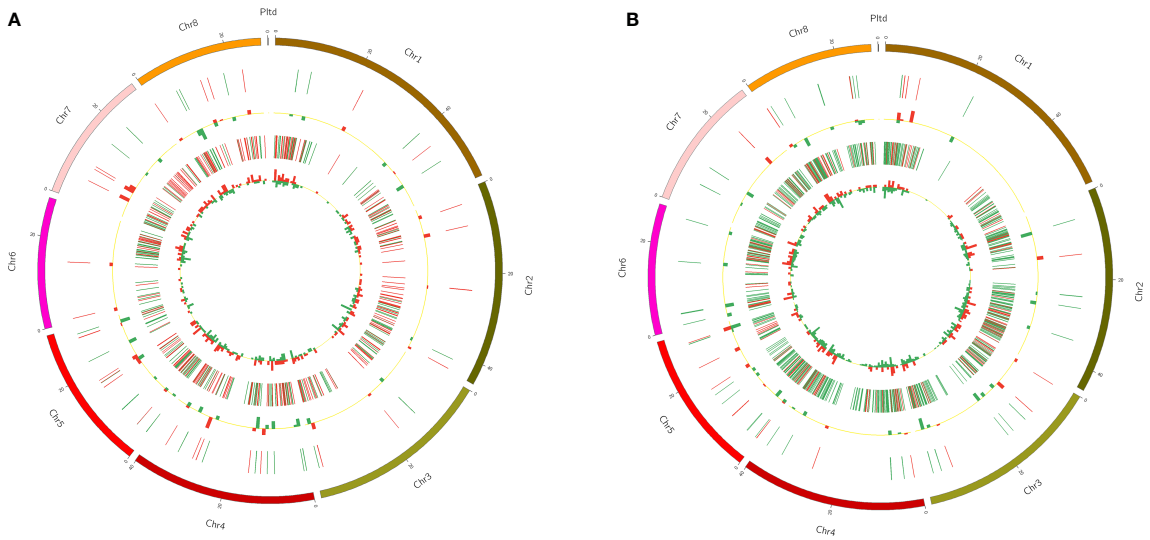


FIGURE 5
The distribution of differentially expressed lncRNAs (DELs) and differentially expressed mRNAs (DEGs) in FC vs FNC **(A)** and FB vs FC **(B)**. Among them, the outermost circle is the distribution of pomegranate chromosomes; the second circle is the distribution of DELs on chromosomes, the red line indicates up-regulation, and the green line indicates down-regulation; the third circle is a histogram of DELs at different positions, and red indicates up-regulation, green means down-regulation. The higher the column, the more the number of DELs; the fourth circle is the distribution of DEGs on chromosomes, the color distribution is the same as that of DELs; the innermost circle is the histogram of DEGs at different positions, the color distribution is same as DELs.

cell wall biogenesis, flavonoid biosynthetic process, cell wall and xyloglucan:xyloglucosyl transferase activity. TCONS_00016729 co-regulated with six DEGs (LOC116187260, LOC116210407, LOC116201343, LOC116202072, LOC116212833 and LOC116202017) participate in secondary metabolic process, chitin catabolic process, cell wall macromolecule catabolic process, polysaccharide catabolic process, lignin catabolic/biosynthetic process, integral component of membrane and cell wall. In addition, TCONS_00017083, TCONS_00018960, TCONS_00022257 and XR_004156511.1 were all engaged in cell wall macromolecule catabolic process, polysaccharide catabolic process, lignin catabolic/biosynthetic process. XR_004155454.1, XR_004157024.1and XR_004157312.1 all engaged in flavonoid biosynthetic process, cell wall biogenesis, xyloglucan metabolic process and xyloglucan:xyloglucosyl transferase activity. As can be seen from the results of FC vs FNC and FB vs FC, phytohormone signaling pathways, lignin catabolic process, lipid transport/binding, cutin biosynthetic process and cell wall played significant roles in fruit non-cracking to fruit cracking. Fruit bagging affected flavonoid biosynthetic process, cell wall biogenesis, xyloglucan metabolic process and xyloglucan:xyloglucosyl transferase activity.

Functional analysis of DEL target genes

Gene regulation may occur in the form of cis-/trans-acting. To explore the putative biological function of DELs, we divined the candidate targets of cis-/trans-acting lncRNA. For the cis-acting lncRNAs, 18 DELs regulate 18 target genes within the upstream and downstream 100 kb range in FC vs FNC (Table S5). For the trans-acting lncRNAs, the results revealed that there were 10 DELs regulated 24 target genes (Table S6; Figure 8A). In FB vs FC, 23 DELs regulate 23 target genes for the cis-acting lncRNAs (Table S7) and 12 DELs regulated 36 target genes (Table S8; Figure 8B). Among them, TCONS_00021929, XR_004158164.1 and TCONS_00001938 have complex regulatory relationships with 17, 1 and 10 target genes, respectively. Besides, four DELs (TCONS_00003922, TCONS_00003921, XR_004154769.1 and XR_004154615.1) have the same target gene LOC116211906. Three DELs (TCONS_00008395, TCONS_00022810, XR_004151985.1) displayed one-to-one associations with three DEGs (LOC116192249, LOC116211189, LOC116201615).

To deeply investigate the function of the *cis*- and *trans*-targets of DELs, GO and KEGG pathways were carried out (Figure 9). GO analysis demonstrated that DELs participated in chitin, calcium ion binding, as well as metal ion binding in FC vs FNC (Figure 9A) and cell wall biogenesis, xyloglucan metabolic process and xyloglucan: xyloglucosyl transferase activity in FB vs FC (Figure 9C). KEGG analysis showed that DELs were involved in the phospholipase D signaling pathway (ko04072) and glycerophospholipid metabolism (ko00564) in FC vs FNC (Figure 9B) and plant hormone signal transduction (ko04075) and starch and sucrose metabolism (ko00500) in FB vs FC (Figure 9D).

10.3389/fpls.2022.943547

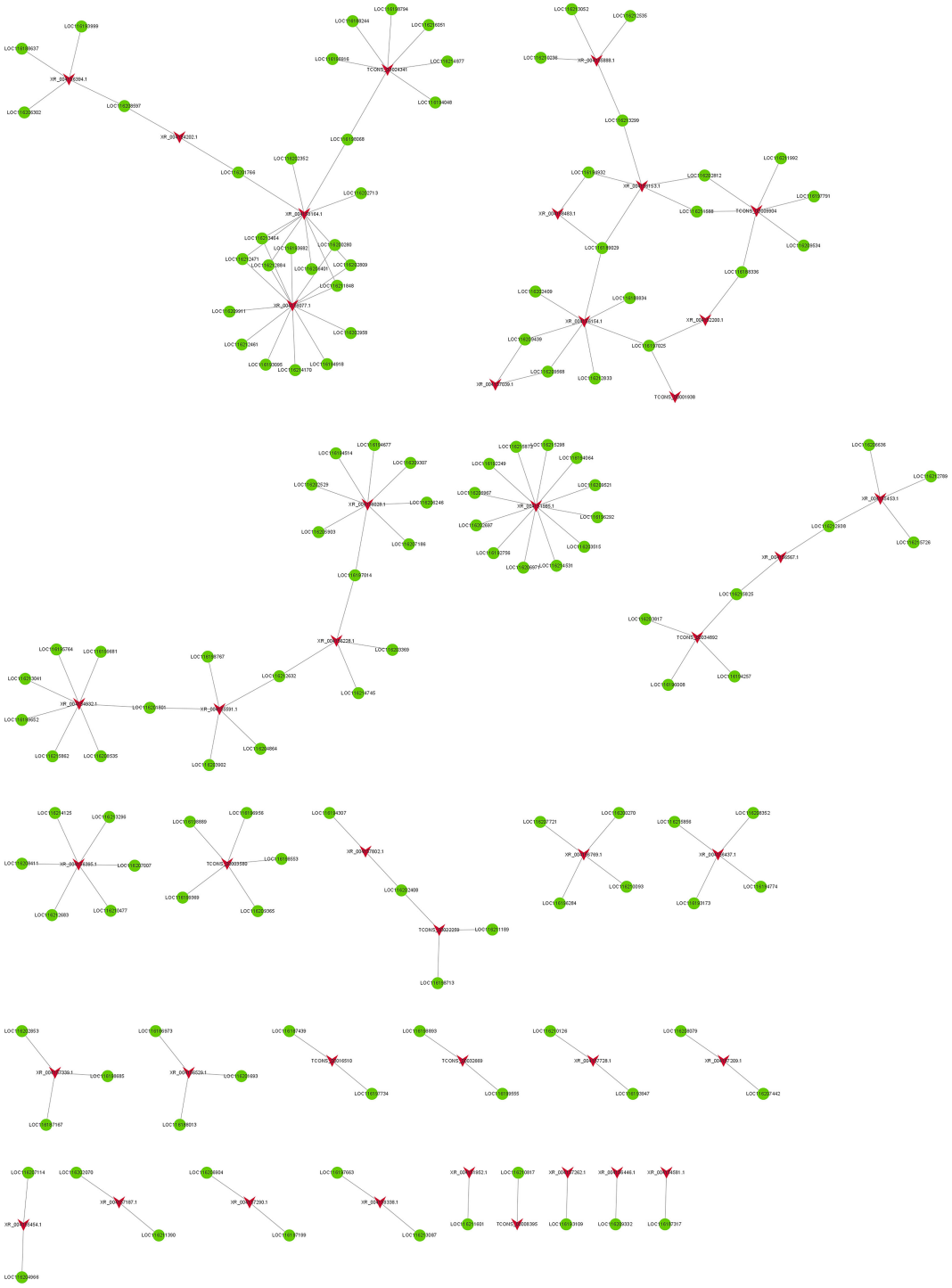


FIGURE 6
The co-expression network of differentially expressed lncRNAs (DELs) and differentially expressed mRNAs (DEGs) in FC vs FNC. The red arrow node and the green circular node represent DELs and DEGs.

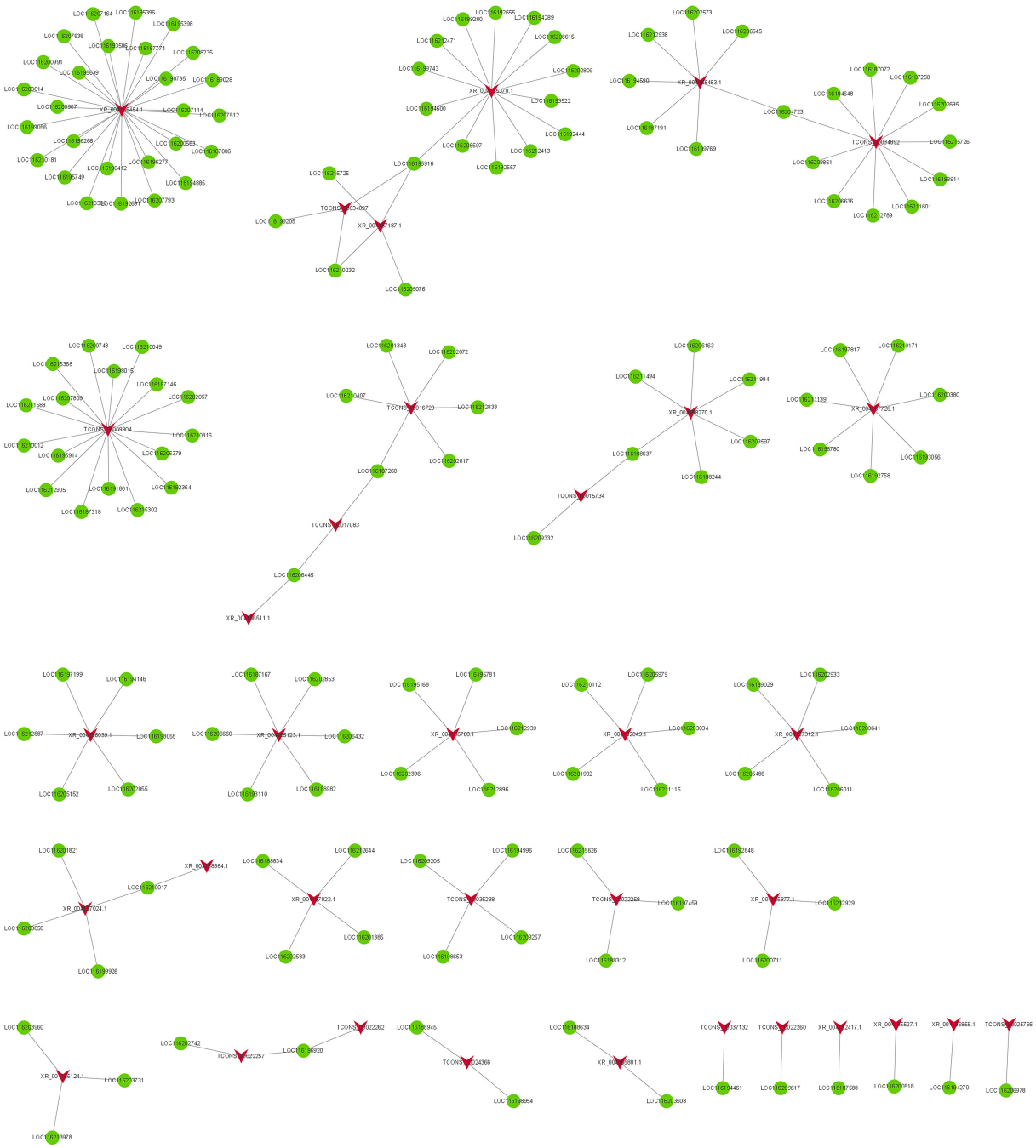


FIGURE 7
The co-expression network of differentially expressed lncRNAs (DELs) and differentially expressed mRNAs (DEGs) in FB vs FC. The red arrow node and the green circular node represent DELs and DEGs.

Quantitative real-time PCR (qRT-PCR) validation of DELs

To validate the reliability and authenticity of the RNA-seq data, ten DELs were randomly selected by quantitative real-time PCR (qRT-PCR) to determine their expression levels. The results showed that DELs displayed similar expression inclination to those obtained by the transcriptome data, which verify the reproducibility and reliability of transcriptomic analysis results (Figure S1). It would be helpful for further studies of the

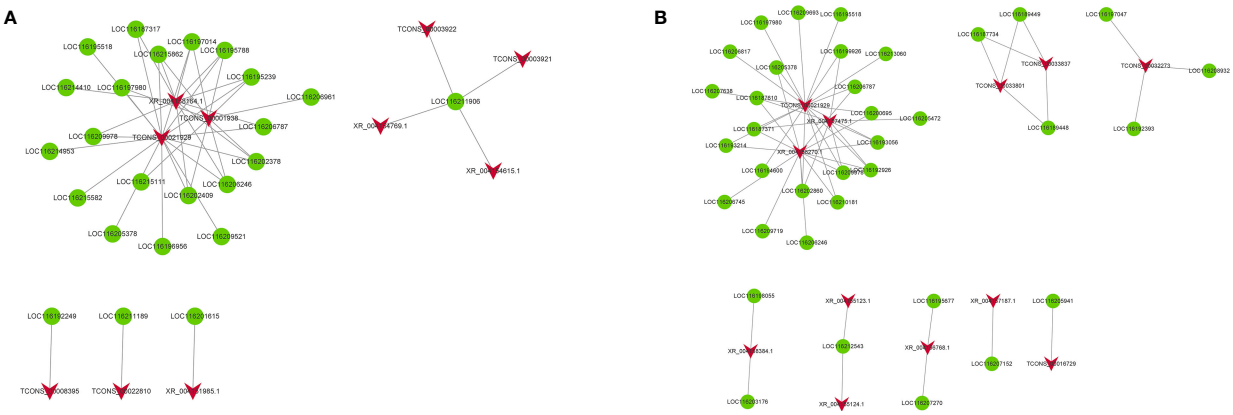


FIGURE 8
The coexpression network of trans-acting differentially expressed lncRNAs (DELs) and differentially expressed mRNAs (DEGs) in FC vs FNC **(A)** and FB vs FC **(B)**. The red arrow node and the green circular node represent DELs and DEGs.

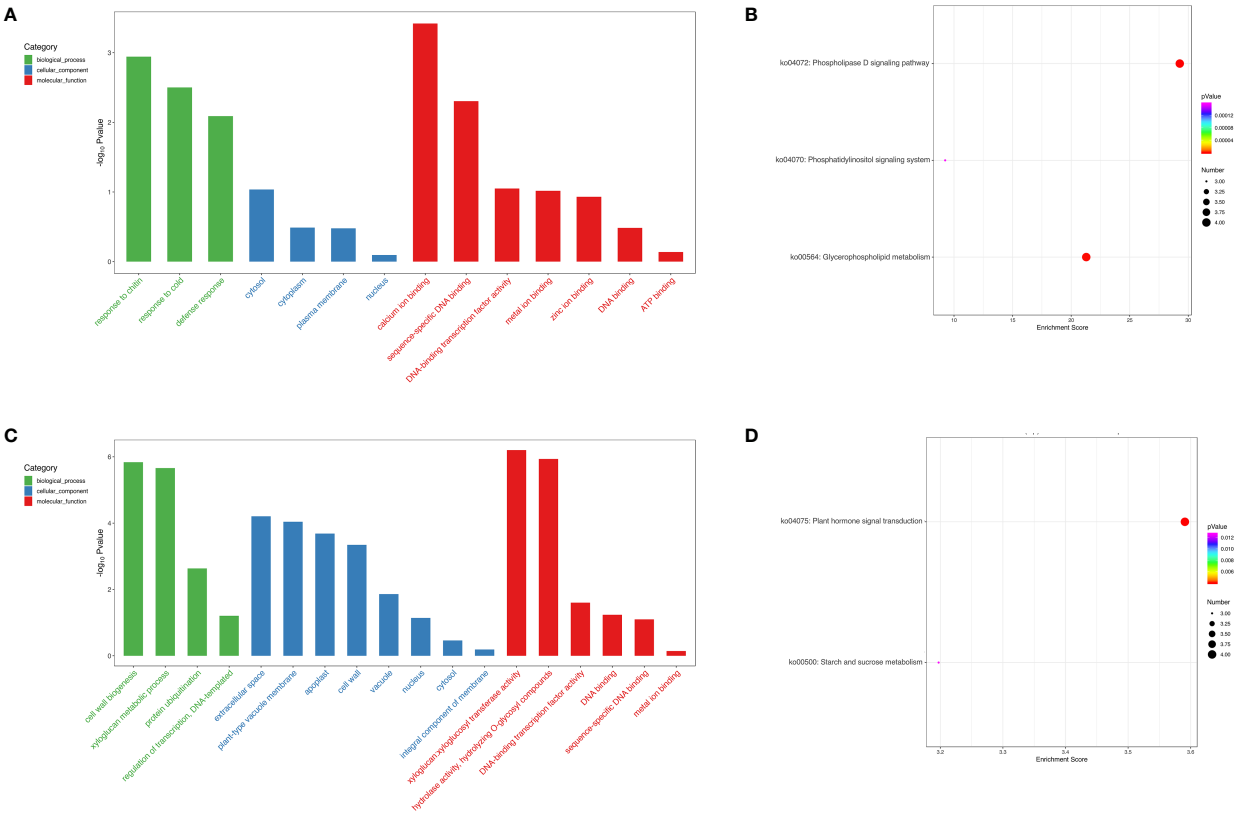


FIGURE 9
Functional Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of long non-coding RNA (lncRNA) target genes. GO enrichment analysis of lncRNA targets genes in FC vs FNC **(A)** and FB vs FC **(C)**. KEGG enrichment analysis of lncRNA targets genes FC vs FNC **(B)** and FB vs FC **(D)**.

lncRNAs that involve in lipid transport, wax biosynthetic process, cell wall organization, phytohormone signaling pathways and calcium ion binding in fruit cracking and fruit non-cracking of pomegranate.

Discussion

Analysis of fruit cracking under bagging

Fruit cracking occurs during fruit development and greatly reduces the quality, shelf life and marketability of the fruit, resulting in tremendous economic losses. Among the factors that cause fruit cracking, in addition to environmental factors, the resistance of the fruit itself to cracking is crucial. Therefore, the mechanism of fruit cracking is still poorly understood.

There are multiple types of fruit cracking, such as longitudinal cracking, transversal cracking, longitudinal and transversal cracking, irregular cracking, calyx-end cracking, internal ring cracking, circular cracking and cuticular cracking (Ginzberg and Stern, 2019; Wang et al., 2021). In the present study, pomegranate fruit cracking types were longitudinal cracking, transversal cracking, longitudinal and transversal cracking, irregular cracking (Figure 1). Among them, the proportion of longitudinal cracking accounted for more than half of the amount. The fruit cracking rate of pomegranate under bagging treatment was less than 10%, while that of the control group was about 40%, indicating that bagging played a gigantic role in reducing the fruit cracking rate, which was consistent with previous studies (Griñán et al., 2019; Asrey et al., 2020).

An dataset of lncRNAs associated with fruit cracking in pomegranate

LncRNAs exert an essential role in diverse biological processes such as leaf-color mutation, fruit development and ripening of plants and have been extensively studied in recent years (Zhang et al., 2018; Wu et al., 2019b). High-throughput sequencing technology for transcriptome analysis has been expeditiously ameliorated and dozens of lncRNAs have been identified in plants. For instance, 25,699, 3857, 6253 and 3679 lncRNAs were identified in Fragaria pentaphylla, Cucumis melo, Pyrus pyrifolia and Solanum lycopersicum, respectively (Zhu et al., 2015; Bai et al., 2019; Tian et al., 2019; Li et al., 2020a). Nevertheless, little is known about lncRNAs in pomegranate fruit cracking. In the present study, 3194 lncRNAs were identified with a total length of 4898846 nt and an average length of 1533.77 nt by strict filtration, among them, 61 DELs in FC vs FNC and 62 DELs in FB vs FC were differentially expressed. LncRNAs expression levels was confirmed by qRT-PCR analysis, which fully demonstrated the authenticity and accuracy of transcriptome results. Furthermore, we revealed that the distribution of DELs and DEGs on chromosomes was not uneven (Figure 3), this was consistent in the case of S. lycopersicum (Zhu et al., 2015).

Functional analysis of DELs and DEGs

To further analyze the functions between DELs and DEGs, we constructed a coexpression network of 42 DELs and 137 DEGs in FC vs FNC (Figure 6) and 35 DELs and 160 DEGs in FB vs FC (Figure 7). The DEL XR_004158164.1 co-regulated with ten DEGs involved in ethylene-activated signaling pathway, lignin catabolic process, lipid transport/binding, cutin biosynthetic process and cell wall in FC vs FNC. The cuticle, composed of cutin and wax, is an essential part of the pericarp and plays a significant role in fruit cracking (Domínguez et al., 2011; Espaa et al., 2014; Petit et al., 2016; Jakobson et al., 2016). Ethylene is extensively involved in plant growth and development, and plays an essential regulatory role in fruit ripening, softening and cracking (Bapat et al., 2010; Chen et al., 2019). ERFs are located downstream of the ethylene signaling pathway, previous studies have shown that ERF4 was associated with fruit ripening and cracking in watermelon and tomato (Liao et al., 2019; Hu et al., 2022).

The main function of polygalacturonase (PG) is to hydrolyze polygalacturic acid, degrade pectin and disintegrate cell wall (Dautt-Castro et al., 2019). There is increasing evidence that PGs play an important role in fruit development in tomato, strawberry, apple, apricot and pear (Chen et al., 2016; Ke et al., 2018; Hou et al., 2019; Zhang et al., 2019; Paniagua et al., 2020). Dautt-Castro et al. (2019) found that MiPG14/17/21/22/49/69 were highly upregulated during fruit ripening. In addition, it was shown that PG was involved in pectin degradation in atemoya cv. African Pride, affecting fruit ripening and dehiscence (Chen et al., 2019). The PG homologs ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1) and ADPG2 were key factors causing fruit pod cracking in Arabidopsis thaliana (Ogawa et al., 2009). In our study, TCONS_00001938 co-regulated with LOC116197025 (Polygalacturonase-1, PG1) involved in lipid transport, wax biosynthetic process and cell wall organization. Zhu et al. (2020) found that genes related to cell wall metabolism and cuticle biosynthesis play a pivotal role in fruit cracking by performing peel transcriptome sequencing of the grape berry cv. Xiangfei (Vitis vinifera L.). In addition, XR_004157290.1 co-regulated with other genes such as LOC116197199 (Polygalacturonase-2, PG2) involved in ethylene-activated/abscisic acid-activated/jasmonic acid mediated signaling pathway to affect fruit cracking.

In FB vs FC, XR_004155454.1, XR_004157024.1and XR_004157312.1 were involved in flavonoid biosynthetic process, cell wall biogenesis, xyloglucan metabolic process and xyloglucan:xyloglucosyl transferase activity.

Cis-/*trans*-acting of the lncRNAs and their target genes involved in lipid metabolism, chitin, and calcium ion binding, cell wall biogenesis, xyloglucan metabolic process

LncRNAs may regulate gene expression in the form of *cis*-/*trans*-acting manner (Fatica and Bozzoni, 2014). Cis-regulation refers to the transcriptional activation and expression regulation of adjacent mRNA by non-coding RNA, while trans, on the contrary, regulates the transcription of distal mRNA (Yan et al., 2017; Wu et al., 2019a). In this study, 18 DELs may regulate 18 target genes and 10 DELs regulated 24 target genes by cis-/trans-acting manner in FC vs FNC (Figure 8A), which participate in lipid metabolism and may be a crucial factor for fruit cracking. In addition, 23 DELs regulate 23 target genes for the cis-acting lncRNAs and 12 DELs regulated 36 target genes in FB vs FC (Figure 8B).

To investigate the function of the *cis*- and *trans*-targets of DELs, GO and KEGG pathways were carried out (Figure 9). GO analysis demonstrated that DELs participated in chitin, calcium ion binding, as well as metal ion binding in FC vs FNC and cell wall biogenesis, xyloglucan metabolic process and xyloglucan: xyloglucosyl transferase activity in FB vs FC. Calcium, an essential mineral element for plant growth and development, can interact with pectic acid to form calcium pectinate, which stabilizes the cell wall structure and strengthens the mechanical strength of the fruit pericarp, thus enhancing the fruit's resistance to cracking (Hernández-Muñoz et al., 2006; Bakeer, 2016). Studies have shown that exogenous sprays of calcium or its compounds can reduce fruit cracking in pomegranates and cherries (Hosein-Beigi et al., 2019; Davarpanah et al., 2018). In addition, some metal ion (zinc and potash) also affect pomegranate cracking (Sharifani, 2014). KEGG analysis showed that DELs were involved in the phospholipase D signaling pathway (ko04072) and glycerophospholipid metabolism (ko00564). Phospholipase D hydrolyzes the phosphodiester bond of glycerol lipid phosphatidylcholine to produce phosphatidylic acid (Jenkins and Frohman, 2005). Glycospholipid metabolism belongs to lipid metabolism (Lu et al., 2022; Li et al., 2022), it suggests that lipid metabolism plays significant role in fruit cracking. In FB vs FC, DELs were participated in plant hormone signal transduction (ko04075) and starch and sucrose metabolism (ko00500), indicating bagging affected hormone biosynthesis, starch and sucrose metabolism, thereby influencing fruit dehiscence.

Conclusion

In conclusion, the incidence of fruit cracking was 7.26% under the bagging treatment, the control group was as high as 38.11%, indicating that bagging substantially reduced the fruit cracking rate. In addition, a total of 3194 lncRNAs were obtained in pomegranate, 61 DELs and 62DELs were identified in FC vs FNC and FB vs FC. We constructed a coexpression network of 42 DELs and 137 DEGs in FC vs FNC and 35 DELs and 160 DEGs in FB vs FC, suggesting that there are involved in phytohormone signaling pathway, lignin catabolic process, lipid transport/binding, cutin biosynthetic process and cell wall organization. Moreover, the results demonstrated that 18 cis-acting DELs regulate 18 target genes, and 10 trans-acting DELs regulate 24 different target genes in FC vs FNC. Among FB vs FC, 23 DELs regulate 23 target genes for the cis-acting lncRNAs and 12 DELs regulated 36 target genes. GO and KEGG analysis results showed that DELs were involved in calcium ion binding, metal ion binding and glycerophospholipid metabolism in FC vs FNC, flavonoid biosynthetic process, cell wall biogenesis, xyloglucan metabolic process, hormone signal transduction and starch and sucrose metabolism in FB vs FC. This study provide insight into the molecular mechanism of fruit cracking in pomegranate, and lay a foundation for future research on fruit cracking of other fruits.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA773212.

Author contributions

Conceptualization, YYW and ZY; Experimental operation, YYW and YZ; Writing—original draft preparation, YYW; Writing—review and editing YYW, YZ, YQW, XZ, ZH, and HL; Project administration, ZY. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Initiative Project for Talents of Nanjing Forestry University (GXL2014070, GXL2018032), the Priority Academic Program Development of Jiangsu High Education Institutions (PAPD), the Natural Science Foundation of Jiangsu Province (BK20180768), China Scholarship Council (CSC, 202108320304), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_1113).

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.943547/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The expression patterns of ten lncRNAs.

References

Anders, S., and Huber, W. (2012). *Differential expression of RNA-seq data at the gene level-the DESeq package* Vol. 10 (Heidelberg, Germany: European Molecular Biology Laboratory (EMBL), f1000research.

Asrey, R., Kumar, K., Sharma, R. R., and Meena, N. K. (2020). Fruit bagging and bag color affects physico-chemical, nutraceutical quality and consumer acceptability of pomegranate (*Punica granatum* l.) arils. *J. Food Sci. Tech.* 57, 1469–1476. doi: 10.1007/s13197-019-04182-x

Baghel, R. S., Keren-Keiserman, A., and Ginzberg, I. (2021). Metabolic changes in pomegranate fruit skin following cold storage promote chilling injury of the peel. *Sci. Rep.* 11, 1–13. doi: 10.1038/s41598-021-88457-4

Bai, L., Chen, Q., Jiang, L., Lin, Y., Ye, Y., Liu, P., et al. (2019). Comparative transcriptome analysis uncovers the regulatory functions of long noncoding RNAs in fruit development and color changes of *Fragaria pentaphylla*. *Hortic. Res.* 6, 15. doi: 10.1038/s41438-019-0128-4

Bakeer, S. M. (2016). Effect of ammonium nitrate fertilizer and calcium chloride foliar spray on fruit cracking and sunburn of manfalouty pomegranate trees. *Sci. Hortic.* 209, 300–308. doi: 10.1016/j.scienta.2016.06.043

Bapat, V. A., Trivedi, P. K., Ghosh, A., Sane, V. A., Ganapathi, T. R., and Nath, P. (2010). Ripening of fleshy fruit: Molecular insight and the role of ethylene. *Biotechnol. Adv.* 28, 94. doi: 10.1016/j.scienta.2016.06.043

Berry, S., and Dean, C. (2015). Environmental perception and epigenetic memory: mechanistic insight through FLC. *Plant J.* 83, 133–148. doi: 10.1111/tpj.12869

Bolger, A. M., Marc, L., and Bjoern, U. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Capel, C., Yuste-Lisbona, F. J., López-Casado, G., Angosto, T., Cuartero, J., Lozano, R., et al. (2017). Multi-environment QTL mapping reveals genetic architecture of fruit cracking in a tomato RIL *Solanum lycopersicum* × *S. pimpinellifolium* population. *Theor. Appl. Genet.* 130, 213–222. doi: 10.1007/s00122-016-2809-9

Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Curr. Opin. Plant Biol.* 27, 207–216. doi: 10.1016/j.pbi.2015.08.003

Chen, J., Duan, Y., Hu, Y., Li, W., and Xie, J. (2019). Transcriptome analysis of atemoya pericarp elucidates the role of polysaccharide metabolism in fruit ripening and cracking after harvest. *BMC Plant Biol.* 19, 1–19. doi: 10.1186/s12870-019-1756-4

Chen, H., Shao, H., Fan, S., Ma, J., Zhang, D., and Han, M. (2016). Identification and phylogenetic analysis of the POLYGALACTURONASE gene family in apple. *Hortic. Plant J.* 2, 241–252. doi: 10.1016/j.hpj.2017.01.004

Correa, J. P. O., Silva, E. M., and Nogueira, F. T. S. (2018). Molecular control by non-coding RNAs during fruit development: from gynoecium patterning to fruit ripening. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01760

Correia, S., Santos, M., Glińska, S., Gapińska, M., Matos, M., Carnide, V., et al. (2020). Effects of exogenous compound sprays on cherry cracking: Skin properties and gene expression. *J. Sci. Food Agr.* 100, 2911–2921. doi: 10.1002/jsfa.10318

Cortés, C., Ayuso, M. C., Palomares, G., Cuartero, J., and Nuez, F. (1983). Relationship between radial and concentric cracking of tomato fruit. *Sci. Hortic.* 21, 323–328. doi: 10.1016/0304-4238(83)90122-X

Dautt-Castro, M., López-Virgen, A. G., Ochoa-Leyva, A., Contreras-Vergara, C. A., and Islas-Osuna, M. A. (2019). Genome-wide identification of mango (*Mangifera indica* l.) polygalacturonases: Expression analysis of family members and total enzyme activity during fruit ripening. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00969

Davarpanah, S., Tehranifar, A., Abadía, J., Val, J., Davarynejad, G., Aran, M., et al. (2018). Foliar calcium fertilization reduces fruit cracking in pomegranate (*Punica granatum* cv. ardestani). *Sci. Hortic.* 230, 86–91. doi: 10.1016/j.scienta.2017.11.023

Domínguez, E., Cuartero, J., and Heredia, A. (2011). An overview on plant cuticle biomechanics. *Plant Sci.* 181, 77–84. doi: 10.1016/j.plantsci.2011.04.016

Espaa, L., Heredia-Guerrero, J., Segado, P., Benítez, J., Heredia, A., and Domínguez, E. (2014). Biomechanical properties of the tomato (*Solanum lycopersicum*) fruit cuticle during development are modulated by changes in the relative amounts of its components. *New Phytol.* 202, 790–802. doi: 10.1111/nph.12727

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi: 10.1038/nrg3606

Galindo, A., Rodríguez, P., Collado-González, J., Cruz, Z. N., Torrecillas, E., Ondoño, S., et al. (2014). Rainfall intensifies fruit peel cracking in water stressed pomegranate trees. *Agr. For. Meteorol.* 194, 29–35. doi: 10.1016/j.agrformet.2014.03.015

Ghosh, S., and Chan, C. K. K. (2016). Analysis of RNA-seq data using TopHat and cufflinks. *Methods Mol. Biol.* 1374, 339–361. doi: 10.1007/978-1-4939-3167-5_18

Ginzberg, I., and Stern, R. A. (2019). Control of fruit cracking by shaping skin traits-apple as a model. *Crit. Rev. Plant Sci.* 38, 1–10. doi: 10.1080/07352689.2019.1698129

Griñán, I., Morales, D., Galindo, A., Torrecillas, A., Pérez-López, D., Moriana, A., et al. (2019). Effect of preharvest fruit bagging on fruit quality characteristics and incidence of fruit physiopathies in fully irrigated and water stressed pomegranate trees. *J. Sci. Food Agr.* 99, 1425–1433. doi: 10.1002/jsfa.9324

Hadjiargyrou, M., and Delihas, N. (2013). The intertwining of transposable elements and non-coding RNAs. *Int. J. Mol. Sci.* 14, 13307–13328. doi: 10.3390/ijms140713307

Hernández-Muñoz, P., Almenar, E., Ocio, M. J., and Gavara., R. (2006). Effect of calcium dips and chitosan coatings on postharvest life of strawberries (*Fragaria* x *ananassa*). *Postharv. Biol. Tec.* 39, 247–253. doi: 10.1016/j.postharvbio.2005.11.006

Hosein-Beigi, M., Zarei, A., Rostaminia, M., and Erfani-Moghadam, J. (2019). Positive effects of foliar application of Ca, b and GA_3 on the qualitative and quantitative traits of pomegranate (*Punica granatum* l.) cv. 'Malase-Torshe-Saveh'. *Sci. Hortic.* 254, 40–47. doi: 10.1016/j.scienta.2019.04.081

Hou, Y., Wu, F., Zhao, Y., Shi, L., and Zhu, X. (2019). Cloning and expression analysis of polygalacturonase and pectin methylesterase genes during softening in apricot (*Prunus armeniaca* l.) fruit. *Sci. Hortic.* 256, 108607. doi: 10.1016/j.scienta.2019.108607

Hu, Y., Sun, H., Han, Z., Wang, S., Wang, T., Li, Q., et al. (2022). ERF4 affects fruit ripening by acting as a JAZ interactor between ethylene and jasmonic acid hormone signaling pathways. *Hortic. Plant J.* 8,1–11 doi: 10.1016/j.hpj.2022.01.002

Jakobson, L., Lindgren, L. O., Verdier, G., Laanemets, K., Brosché, M., Beisson, F., et al. (2016). BODYGUARD is required for the biosynthesis of cutin in arabidopsis. *New Phytol.* 211, 614–626. doi: 10.1111/nph.13924

Jenkins, G. M., and Frohman, M. A. (2005). Phospholipase d: a lipid centric review. *Cell. Mol. Life Sci.* 62, 2305–2316. doi: 10.1007/s00018-005-5195-z

Jiang, F., Lopez, A., Jeon, S., De Freitas, S. T., Yu, Q., Wu, Z., et al. (2019). Disassembly of the fruit cell wall by the ripening-associated polygalacturonase and expansin influences tomato cracking. *Hortic. Res.* 6, 17. doi: 10.1038/s41438-018-0105-3

Joshi, M., Baghel, R. S., Fogelman, E., Stern, R. A., and Ginzberg, I. (2018). Identification of candidate genes mediating apple fruit-cracking resistance following the application of gibberellic acids 4 + 7 and the cytokinin 6-benzyladenine. *Plant Physiol. Bioch.* 127, 436–445. doi: 10.1016/j.plaphy.2018.04.015

Joshi, M., Schmilovitch, Z., and Ginzberg, I. (2021). Pomegranate fruit growth and skin characteristics in hot and dry climate. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.725479

Kang, Y. J., Yang, D. C., Kong, L., Hou, M., Meng, Y. Q., Wei, L., et al. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* 45, W12–W16. doi: 10.1093/nar/gkx428

Ke, X., Wang, H., Li, Y., Zhu, B., Zang, Y., He, Y., et al. (2018). Genome-wide identification and analysis of polygalacturonase genes in *Solanum lycopersicum*. *Int. J. Mol. Sci.* 19 2290, 1–8. doi: 10.3390/ijms19082290

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Kim, E. D., and Sung, S. (2012). Long noncoding RNA: unveiling hidden layer of gene regulatory networks. *Trends Plant Sci.* 17, 16–21. doi: 10.1016/j.tplants.2011.10.008

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Laurent, G. S., Wahlestedt, C., and Kapranov, P. (2015). The landscape of long noncoding RNA classification. *Trends Genet.* 31, 239–251. doi: 10.1016/j.tig.2015.03.007

Lewandowska, M., Keyl, A., and Feussner, I. (2020). Wax biosynthesis in response to danger: its regulation upon abiotic and biotic stress. *New Phytol.* 227, 698–713. doi: 10.1111/nph.16571

Liao, N., Hu, Z., Li, Y., Hao, J., and Zhang, M. (2019). Ethylene-responsive factor 4 is associated with the desirable rind hardness trait conferring cracking resistance in fresh fruits of watermelon. *Plant Biotechnol. J.* 18, 1066–1077. doi: 10.1111/pbi.13276

Li, Q., Cheng, C., Zhang, X., Wang, C., and Yang, S. (2020c). Preharvest bagging and postharvest calcium treatment affects superficial scald incidence and calcium nutrition during storage of 'Chili' pear (*Pyrus bretschneideri*) fruit. *Postharv. Biol. Tec.* 163, 111149. doi: 10.1016/j.postharvbio.2020.111149

Li, N., Fu, L., Song, Y., Li, J., Xue, X., Li, S., et al. (2020b). Wax composition and concentration in jujube (Ziziphus jujuba mill.) cultivars with differential resistance to fruit cracking. *J. Plant Physiol.* 255, 153294. doi: 10.1016/j.jplph.2020.153294

Li, Z., Hu, J., Wu, Y., Wang, J., Song, H., Chai, M., et al. (2022). Integrative analysis of the metabolome and transcriptome reveal the phosphate deficiency response pathways of alfalfa. *Plant Physiol. Bioch.* 170, 49–63. doi: 10.1016/j.plaphy.2021.11.039

Li, L., Liu, J., Liang, Q., Zhang, Y., and Li, Y. (2020a). Genome-wide analysis of long noncoding RNAs affecting floral bud dormancy in pears in response to cold stress. *Tree Physiol.* 41, 771–790. doi: 10.1093/treephys/tpaa147

Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., et al. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in arabidopsis. *Plant Cell* 24, 4333–4345. doi: 10.1105/tpc.112.102855

Liu, C., Zhao, Y., Zhao, X., Dong, J., and Yuan, Z. (2020). Genome-wide identification and expression analysis of the CLC gene family in pomegranate (*Punica granatum*) reveals its roles in salt resistance. *BMC Plant Biol.* 20, 560. doi: 10.21203/rs.3.rs-54027/v1

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

Li, L., Wang, X., Rajkumar, S., Viktor, S., Wei, D., Hang, H., et al. (2007). Global identification and characterization of transcriptionally active regions in the rice genome. *PLoS One* 2, e294. doi: 10.1371/journal.pone.0000294

Li, F., Zhang, X., Wang, J., Jiang, Y., Zhang, X., and Li, X. (2021). Preharvest application of 1-methylcyclopropene and ethephon altered cuticular wax biosynthesis and fruit quality of apples at harvest and during cold storage. *Hortic. Plant J.* 8, 143–152. doi: 10.1016/j.hpj.2021.11.008

Li, A., Zhang, J., and Zhou, Z. (2014). PLEK: A tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinf.* 15, 1–10. doi: 10.1186/1471-2105-15-311

Lu, M., Li, Y., Jia, H., Xi, Z., Gao, Q., Zhang, Z. Z., et al. (2022). Integrated proteomics and transcriptome analysis reveal a decreased catechins metabolism in variegated tea leaves. *Sci. Hortic.* 295, 110824. doi: 10.1016/j.scienta.2021.110824

Lyu, Y., Porat, R., Yermiyahu, U., Heler, Y., Holland, D., and Dag, A. (2020). Effects of nitrogen fertilization on pomegranate fruit, aril and juice quality. *J. Sci. Food Agric.* 100, 1678–1686. doi: 10.1002/jsfa.10182

Niu, J., Shi, Y., Huang, K., Zhong, Y., Chen, J., Sun, Z., et al. (2020). Integrative transcriptome and proteome analyses provide new insights into different stages of *Akebia trifoliata* fruit cracking during ripening. *Biotechnol. Biofuels* 13, 149. doi: 10.1186/s13068-020-01789-7

Ogawa, M., Kay, P., Wilson, S., and Swain, S. M. (2009). ARABIDOPSIS DEHISCENCE ZONE POLYGALACTURONASE1 (ADPG1), ADPG2, and QUARTET2 are polygalacturonases required for cell separation during reproductive development in arabidopsis. *Plant Cell* 21, 216–233. doi: 10.1105/tpc.108.063768

Ono, N. N., Britton, M. T., Fass, J. N., Nicolet, C. M., Lin, D., and Tian, L. (2011). Exploring the transcriptome landscape of pomegranate fruit peel for natural product biosynthetic gene and SSR marker discovery. *J. Integr. Plant Biol.* 53, 800–813. doi: 10.1111/j.1744-7909.2011.01073.x

Paniagua, C., Ric-Varas, P., García-Gago, J. A., López-Casado, G., Blanco-Portales, R., Muñoz-Blanco, J., et al. (2020). Elucidating the role of polygalacturonase genes in strawberry fruit softening. *J. Exp. Bot.* 71, 7103–7117. doi: 10.1093/jxb/eraa398

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Petit, J., Bres, C., Mauxion, J. P., Fabienne, W. J. T., Martin, L. B. B., Fich, E. A., et al. (2016). The glycerol-3-Phosphate acyltransferase GPAT6 from tomato plays a central role in fruit cutin biosynthesis. *Plant Physiol.* 171, 894–913. doi: 10.1104/pp.16.00409

Qin, G., Liu, C., Li, J., Qi, Y., Gao, Z., Zhang, X., et al. (2020). Diversity of metabolite accumulation patterns in inner and outer seed coats of pomegranate: exploring their relationship with genetic mechanisms of seed coat development. *Hortic. Res.* 7, 10. doi: 10.1038/s41438-019-0233-4

Qin, G., Xu, C., Ming, R., Tang, H., Guyot, R., Kramer, E. M., et al. (2017). The pomegranate (*Punica granatum* l.) genome and the genomics of punicalagin biosynthesis. *Plant J.* 91, 1108–1128. doi: 10.1111/tpj.13625

Quero-García, J., Letourmy, P., Campoy, J. A., Branchereau, C., Malchev, S., Barreneche, T., et al. (2021). Multi-year analyses on three populations reveal the first stable QTLs for tolerance to rain-induced fruit cracking in sweet cherry (*Prunus avium* l.). *Hortic. Res.* 8, 136. doi: 10.1038/s41438-021-00571-6

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J. L., and Pachter, L. (2011). Improving RNA-seq expression estimates by correcting for fragment bias. *Genome Biol.* 12, 1–14. doi: 10.1186/gb-2011-12-3-r22

Saminathan, T., Bodunrin, A., Singh, N. V., Devarajan, R., Nimmakayala, P., Jeff, M., et al. (2016). Genome-wide identification of microRNAs in pomegranate (*Punica granatum* l.) by high-throughput sequencing. *BMC Plant Biol.* 16, 1–16. doi: 10.1186/s12870-016-0807-3

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Sharifani, M. (2014). Description of biomechanical forces and physiological parameters of fruit cracking in pomegranate. *Sci. Hortic.* 178, 224. doi: 10.1016/j.scienta.2014.09.005

Singh, V., Dan, G., Parimi, P., Kochanek, B., and Friedman, H. (2021). Postharvest calcium treatment of apple fruit increased lenticel breakdown and altered cuticle structure. *Postharv. Biol. Tec.* 171, 111331. doi: 10.1016/j.postharvbio.2020.111331

Sonnhammer, E., Eddy, S. R., Ewan, B., Alex, B., and Richard, D. (1998). Pfam: Multiple sequence alignments and HMM-profiles of protein domains. *Nucleic Acids Res.* 26, 320–322. doi: 10.1093/nar/26.1.320

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41, e166. doi: 10.1093/nar/gkt646

Tian, Y., Bai, S., Dang, Z., Hao, J., Zhang, J., and Hasi, A. (2019). Genome-wide identification and characterization of long non-coding RNAs involved in fruit ripening and the climacteric in *Cucumis melo*. *BMC Plant Biol.* 19, 369. doi: 10.1186/s12870-019-1942-4

Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: Genomics, evolution, and mechanisms. *Cell* 154, 26–46. doi: 10.1016/j.cell.2013.06.020

Vlachojannis, C., Erne, P., Schoenenberger, A. W., and Chrubasik-Hausmann, S. (2015). A critical evaluation of the clinical evidence for pomegranate preparations in the prevention and treatment of cardiovascular diseases. *Phytother. Res.* 29, 501–508. doi: 10.1002/ptr.5280

Wang, S. Y., and Camp, M. J. (2000). Temperatures after bloom affect plant growth and fruit quality of strawberry. *Sci. Hortic.* 85, 183–199. doi: 10.1016/S0304-4238(99)00143-0

Wang, Y., Fan, X., Lin, F., He, G., Terzaghi, W., and Zhu, D. (2014). Arabidopsis noncoding RNA mediates control of photomorphogenesis by red light. *P. Nati. Acad. Sci. U.S.A.* 111, 10359–10364. doi: 10.1073/pnas.1409457111

Wang, J., Gao, X., Ma, Z., Chen, J., and Liu, Y. (2019). Analysis of the molecular basis of fruit cracking susceptibility in *Litchi chinensis* cv. baitangying by transcriptome and quantitative proteome profiling. *J. Plant Physiol.* 234, 106–116. doi: 10.1016/j.jplph.2019.01.014

Wang, Y., Guo, L., Zhao, X., Zhao, Y., Hao, Z., Luo, H., et al. (2021). Advances in mechanisms and omics pertaining to fruit cracking in horticultural plants. *Agronomy* 11, 1045. doi: 10.3390/agronomy11061045

Wang, T. Z., Liu, M., Zhao, M. G., Chen, R., and Zhang, W. H. (2015). Identification and characterization of long non-coding RNAs involved in osmotic and salt stress in *Medicago truncatula* using genome-wide high-throughput sequencing. *BMC Plant Biol.* 15, 131. doi: 10.1186/s12870-015-0530-5

Wang, Y., and Long, L. E. (2015). Physiological and biochemical changes relating to postharvest splitting of sweet cherries affected by calcium application in hydrocooling water. *Food Chem.* 181, 241–247. doi: 10.1016/j.foodchem.2015.02.100

Wang, Y., Yang, X., Chen, Z., Zhang, J., Si, K., Xu, R., et al. (2022). Function and transcriptional regulation of CsKCS20 in the elongation of very-long-chain fatty acids and wax biosynthesis in *Citrus sinensis* flavedo. *Hortic. Res.* 9, uhab027. doi: 10.1093/hr/uhab027

Wu, Y., Guo, J., Wang, T., Cao, F., and Wang, G. (2019b). Transcriptional profiling of long noncoding RNAs associated with leaf-color mutation in *Ginkgo biloba* l. *BMC Plant Biol.* 19, 1–13. doi: 10.1186/s12870-019-2141-z

Wu, X., Shi, T., Iqbal, S., Zhang, Y., Liu, L., and Gao, Z. (2019a). Genome-wide discovery and characterization of flower development related long non-coding RNAs in *Prunus mume*. *BMC Plant Biol.* 19, 1–17. doi: 10.1186/s12870-019-1672-7

Xue, L., Sun, M., Wu, Z., Yu, L., Yu, Q., Tang, Y., et al. (2020). LncRNA regulates tomato fruit cracking by coordinating gene expression *via* a hormone-redox-cell wall network. *BMC Plant Biol.* 20, 162. doi: 10.1186/s12870-020-02373-9

Yan, P., Luo, S., Lu, J. Y., and Shen, X. (2017). Cis- and trans-acting lncRNAs in pluripotency and reprogramming. *Curr. Opin. Genet. Dev.* 46, 170–178. doi: 10.1016/j.gde.2017.07.009

Yuan, Z., Fang, Y., Zhang, T., Fei, Z., Han, F., Liu, C., et al. (2017). The pomegranate (*Punica granatum* l.) genome provides insights into fruit quality and ovule developmental biology. *Plant Biotechnol. J.* 16, 1363–1374. doi: 10.1111/pbi.12875

Zhang, G., Chen, D., Zhang, T., Duan, A., Zhang, J., and He, C. (2018). Transcriptomic and functional analyses unveil the role of long non-coding RNAs in anthocyanin biosynthesis during sea buckthorn fruit ripening. *DNA Res.* 25, 465–476. doi: 10.1093/dnares/dsy017

Zhang, C., Cui, L., Zhang, P., Dong, T., and Fang, J. (2021). Transcriptome and metabolite profiling reveal that spraying calcium fertilizer reduces grape berry cracking by modulating the flavonoid biosynthetic metabolic pathway. *Food Chem.: Mol. Sci.* 2, 100025. doi: 10.1016/j.fochms.2021.100025

Zhang, S., Ma, M., Zhang, H., Zhang, S., Qian, M., Zhang, Z., et al. (2019). Genome-wide analysis of polygalacturonase gene family from pear genome and identification of the member involved in pear softening. *BMC Plant Biol.* 19, 1–12. doi: 10.1186/s12870-019-2168-1

Zhao, Y., Liu, C., Ge, D., Yan, M., and Yuan, Z. (2020). Genome-wide identification and expression of YABBY genes family during flower development in *Punica granatum* l. *Gene* 752, 144784. doi: 10.1016/j.gene.2020.144784

Zhao, X., Yuan, Z., Feng, L., and Fang, Y. (2015). Cloning and expression of anthocyanin biosynthetic genes in red and white pomegranate. *J. Plant Res.* 128, 687–696. doi: 10.1007/s10265-015-0717-8

Zhu, Q.H., Stephen, S., Taylor, J., Helliwell, C. A., and Wang, M. B. (2014). Long noncoding RNA s responsive to fusarium oxysporum infection in *Arabidopsis thaliana*. *New Phytol.* 201, 574–584. doi: 10.1111/nph.12537

Zhu, B., Yang, Y., Li, R., Fu, D., Wen, L., Luo, Y., et al. (2015). RNA Sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. *J. Exp. Bot.* 66, 4483–4495. doi: 10.1093/jxb/erv203

Zhu, M., Yu, J., Zhao, M., Wang, M., and Yang, G. (2020). Transcriptome analysis of metabolisms related to fruit cracking during ripening of a cracking-susceptible grape berry cv. xiangfei (*Vitis vinifera* l.). *Genes Genom.* 42, 639–650. doi: 10.1007/s13258-020-00930-y

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 November 2022
DOI 10.3389/fpls.2022.1024515

OPEN ACCESS

EDITED BY
Shunquan Lin,
South China Agricultural University, China

REVIEWED BY
Tangchun Zheng,
Beijing Forestry University, China
Yuanyuan Jiang,
Shaoguan University, China

*CORRESPONDENCE
Danlong Jing
jingdanlong110@126.com
Qigao Guo
qgguo@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 21 August 2022
ACCEPTED 19 October 2022
PUBLISHED 03 November 2022

CITATION
Yu Y, Yang M, Liu X, Xia Y, Hu R, Xia Q, Jing D and Guo Q (2022) Genome-wide analysis of the *WOX* gene family and the role of *EjWUSa* in regulating flowering in loquat (*Eriobotrya japonica*). *Front. Plant Sci.* 13:1024515. doi: 10.3389/fpls.2022.1024515

Genome-wide analysis of the *WOX* gene family and the role of *EjWUSa* in regulating flowering in loquat (*Eriobotrya japonica*)

Yuanhui Yu[1,2†], Miaomiao Yang[1,2†], Xinya Liu[1,2], Yan Xia[1,2], Ruoqian Hu[1,2], Qingqing Xia[1,2], Danlong Jing[1,2*] and Qigao Guo[1,2*]

[1]Key Laboratory of Horticulture Science for Southern Mountains Regions of Ministry of Education, College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China, [2]State Cultivation Base of Crop Stress Biology for Southern Mountainous Land of Southwest University, Academy of Agricultural Sciences, Southwest University, Chongqing, China

The WUSCHEL (WUS)-related homeobox (*WOX*) gene family plays a crucial role in stem cell maintenance, apical meristem formation, embryonic development, and various other developmental processes. However, the identification and function of *WOX* genes have not been reported in perennial loquat. In this study, 18 *EjWOX* genes were identified in the loquat genome. Chromosomal localization analysis showed that 18 *EjWOX* genes were located on 12 of 17 chromosomes. Gene structure analysis showed that all *EjWOX* genes contain introns, of which 11 *EjWOX* genes contain untranslated regions. There are 8 pairs of segmental duplication genes and 0 pairs of tandem duplication genes in the loquat *WOX* family, suggesting that segmental duplications might be the main reason for the expansion of the loquat *WOX* family. A *WOX* transcription factor gene named *EjWUSa* was isolated from loquat. The EjWUSa protein was localized in the nucleus. Protein interactions between EjWUSa with EjWUSa and EjSTM were verified. Compared with wild-type *Arabidopsis thaliana*, the *35S::EjWUSa* transgenic *Arabidopsis* showed early flowering. Our study provides an important basis for further research on the function of *EjWOX* genes and facilitates the molecular breeding of loquat early-flowering varieties.

KEYWORDS

loquat, *WOX* gene family, WUS transcription factor, flowering time, protein interaction

Introduction

Loquat is a subtropical evergreen fruit tree of the Rosaceae family (Cao et al., 2013; Zhang et al., 2016). Compared with many fruit trees, loquat has unique characteristics of flowering in winter and fruit ripening in late spring or early summer (Jiang et al., 2019a). Previous studies have found that *EjTFL1*, *EjFRI*, *EjFT*, *EjGI*, *EjCO*, *EjFD*, *EjSOC1*, *EjLEY*, *EjSVP*, *EjAP1*, and *EjSPL* played important roles in regulating loquat flowering (Chen et al., 2020; Esumi et al., 2005; Jiang et al., 2019a, 2019b, 2019c; Reig et al., 2017; Liu et al., 2013; Zhang et al., 2019a; 2016b). However, studies of transcription factors regulating loquat flowering were still limited.

Plant flowering is regulated by both external environmental factors and internal genetic factors. At present, the understanding of angiosperm flowering relies mainly on the studies of flowering regulation in the model plant *Arabidopsis*. In *Arabidopsis*, flower bud differentiation is mainly regulated by the photoperiod pathway, vernalization pathway, gibberellin (GA) pathway, autonomous flowering pathway, heat-sensing pathway, and age pathway. In this regulatory network, about 180 genes interact to regulate *Arabidopsis* flowering (Bergonzi et al., 2013; Zhou et al., 2013). Compared with *Arabidopsis*, the studies of flowering regulation in woody plants are relatively lacking and need further research.

The *WOX* family is a class of plant-specific transcription factors (Feng et al., 2021). Its members possess 60-65 amino acid residues with the helix-loop-helix-turn-helix domain (referred to as homeodomain) and specifically bind DNA by the homeodomain to activate or depress the expression of the target gene in plants (Ikeda et al., 2009; Shafique Khan et al., 2021). *WOX* genes were divided into three separate clades, modern/WUS clade (WC), intermediate clade (IC), and ancient clade (AC) according to the time of their appearance during plant evolution (Alvarez et al., 2018). Based on phylogenetic analysis, *WOX* genes were further divided into nine subgroups (Zhang et al., 2010). In addition to the homeodomain, some WOX proteins contain three other functional domains: the acidic region (Rich in glutamic acid and aspartic acid), the WUS-box (T-L-X-L-F-P-X-X, X is an uncertain amino acid), and the EAR-like motif (X-L-X-L-X-L, X is an uncertain amino acid). The WUS-box is critical for regulating stem cell identity and floral meristem size. The acidic region is the only activation domain of the WUS proteins. In addition to the WUS-box, the EAR-like motif is also a repression domain (Ikeda et al., 2009). The *WUS* gene is the earliest gene identified in the *WOX* gene family (Xu, 2021). The negative feedback loop between WUS and CLAVATA3 (CLV3) underlies the maintenance of stem cell homeostasis in the shoot apical meristem (SAM) (Yadav, 2012; Xiao et al., 2018; Lopes et al., 2021). Previous studies have demonstrated that WUS protein regulates the expression of *CLV3* gene in the organizing center and central zone by forming homodimers with itself and heterodimers with SHOOT MERISTEMLESS (STM), respectively. In turn, CLV3 forms a signaling complex with CLAVATA1 (CLV1) and CLAVATA2 (CLV2) to regulate the expression of the *WUS* gene in the organizing center (Daum et al., 2014; Perales et al., 2016; Zhou et al., 2018; Su et al., 2020). Previous studies have demonstrated that the *WOX* genes regulated plant flowering and development (Tvorogova et al., 2021). However, genome-wide identification and functional analysis of the *WOX* genes in loquat have not been reported.

In this study, we systematically identified the *WOX* family in the loquat genome and analyzed the chromosomal localization, gene structure, conserved motifs, and basic characteristics of the loquat *WOX* family. A *WOX* transcription factor gene named *EjWUSa* was isolated from the triploid loquat 'Wuhezaoyu'. The subcellular localization of EjWUSa was observed in *Nicotiana benthamiana* leaves. Meanwhile, the interactions between EjWUSa with EjWUSa and EjSTM were verified. Finally, we overexpressed *EjWUSa* in wild-type *Arabidopsis* for functional analysis.

Materials and methods

Plant materials and growth conditions

Triploid loquat 'Wuhezaoyu' was cultivated in the orchard of loquat resources belonging to the College of Horticulture and Landscape Architecture, Southwest University. In our previous study, the development of loquat buds was divided into nine stages (Jing et al., 2020), and from July to November 2021, the loquat flower buds at 9 stages were collected from the 12 years of triploid loquat 'Wuhezaoyu'. After removing the superficial fluff, the loquat flower buds were immediately frozen in liquid nitrogen and stored in an ultra-low temperature refrigerator at −80°C until use.

Wild-type *Arabidopsis* was used for stable transformation of the *EjWUSa* gene. Tobacco was used in transient expression assays. Both *Arabidopsis* and tobacco were grown under long-day conditions (16 h light/8 h dark) at 22°C in a controlled environment room. *Arabidopsis* leaves for qRT-PCR were immediately frozen in liquid nitrogen and stored in an ultra-low temperature refrigerator at −80°C until use.

Identification of the *EjWOX* genes in loquat

The Hidden Markov Model (HMM) of the Homeobox (*HOX)* superfamily (PF00046) was obtained at the Pfam website (http://pfam.xfam.org/). The WOX protein sequences in the model plant *Arabidopsis* were obtained from the TAIR

database (https://www.arabidopsis.org/). The *WOX* family is one of six families in the *HOX* superfamily (Feng et al., 2021). The HMM of *HOX* superfamily was used as a query to search for candidate *EjHOX* genes in the loquat genome (The data presented in the study are deposited in NGDC repository accession No. GWHBOTF00000000). A total of 134 candidate *EjHOX* genes were identified in loquat. The conserved domain of candidate *EjHOX* genes was analyzed in NCBI (http://www.ncbi.nlm.nih.gov). As shown in Figure S1, 18 *EjHOX* genes and 15 *AtWOX* genes were in the same branch. Therefore, 18 *EjWOX* genes were finally identified in loquat and named according to their homology with the *AtWOX* genes (Figure S1; Table S1).

Phylogenetic tree, multiple sequence alignment, and characterizations analysis of the WOX proteins

The apple genome file (V 3.0) was downloaded from the apple information resource (GDR, https://www.rosaceae.org) (Wang, 2021). Similar to the *EjHOX* genes, 130 candidate *MdHOX* genes were identified in the apple genome. The conserved domain of candidate *MdHOX* genes was analyzed in NCBI. As shown in Figure S2, 17 *MdHOX* genes and 15 *AtWOX* genes were in the same branch. Therefore, 17 *MdWOX* genes were finally identified in apple and named according to their homology with the *AtWOX* genes (Figure S2; Table S2).

The WOX protein sequences in the model plant tomato were obtained from the article published by Li et al. (2018). A phylogenetic tree including WOX proteins in loquat, apple, *Arabidopsis*, and tomato was established by MEGA software (v 11.0) with the Neighbor-joining method based on the following parameters: pairwise deletion and bootstrap analysis with 1000 replicates (Munir et al., 2016). The phylogenetic tree was imported into the iTOL website (https://itol.embl.de/) for further beautification. Multiple sequence alignment of WOX proteins in loquat were analyzed using DNAMAN software (v 9.0) with the default settings. The ExPASY database (https://web.expasy.org/protparam/) was used to forecast the characteristics of EjWOX proteins including the coding sequence length, theoretical isoelectric point (PI), molecular weight, and amino acid length (Wilkins et al., 1999).

Chromosomal localization, gene structure, conserved motif, and synteny analysis of the *EjWOX* genes in loquat

The MEME suite (http://meme-suite.org/tools/meme) was used to identify the conserved motifs (Bailey et al., 2009). Chromosomal localization, gene structure, conserved motif, and synteny analysis of the *EjWOX* genes were visualized using TBtools software (v 1.098763) (Chen et al., 2020).

Cloning and sequence analysis of genes and promoter

Genes and promoter sequences were obtained in the loquat genome. Gene was amplified from the cDNA of the triploid loquat 'Wuhezaoyu' flower buds by PCR using Phanta Max Super-Fidelity DNA polymerase (Vazyme, China). The PCR product was ligated with the pMD19-T vector. Then, the cloned product was sequenced. Finally, multiple sequence alignment was performed using DNAMAN software. Based on the same method, the promoter was isolated from the DNA of 'Wuhezaoyu'. The specific primers were listed in Table S3. Then, the putative cis-acing elements on the promoter region were found in the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Wilkins et al., 2005). The result was visualized using TBtools software.

Gene expression analysis with qRT-PCR

The total RNA was extracted by EASYspin Plus plant RNA extraction kit (Aidlab, China), and the cDNA was synthesized using PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa, Japan). qRT-PCR was performed using NovoStart® SYBR qPCR SuperMix plus (Novoprotein, China). The loquat *EjActin* gene and *Arabidopsis AtActin* gene were used as internal controls, with the special primers in Table S3. Three biological replicates were applied and data were analyzed with the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Subcellular localization of EjWUSa

The coding sequence (without stop codon) of *EjWUSa* was cloned into the modified pCAMBIA1300 vector (Jing et al., 2020) with the special primers in Table S3. The constructed fusion vector or empty vector was transformed into *Agrobacterium strain* GV3101, respectively, and then tobacco leaves were used for transient expression. An Olympus (BX35) fluorescence fluorescence signals.

Bimolecular fluorescence complementation assay

The coding sequence (without stop codon) of *EjWUSa* was constructed into the pXY104 vector and the special primers with restriction sites (*Sal* I and *Bam*H I) were listed in Table S3. The

coding sequence of *EjWUSa* or *EjSTM* was constructed into the pXY106 vector and the special primers with restriction sites (*Sal* I and *Bam*H I) were listed in Table S3. The constructed fusion vectors or empty vectors were transformed into *Agrobacterium strain* GV3101, respectively, and then tobacco leaves were used for transient expression. An Olympus (BX35) fluorescence microscope (Tokyo, Japan) was used to observe the fluorescence signals. Three independent leaves were observed. The method was described by Liu et al. (2021).

Firefly luciferase complementation imaging assay

The coding sequence of *EjWUSa* was constructed into the pCAMBIA-CLuc vector and the special primers with restriction sites (*Sal* I and *Bam*H I) were listed in Table S3. The coding sequence (without stop codon) of *EjWUSa* or *EjSTM* was constructed into the pCAMBIA-NLuc vector and the special primers with restriction sites (*Sal* I and *Bam*H I) were listed in Table S3. The constructed fusion vectors or empty vectors were transformed into *Agrobacterium strain* GV3101, respectively, and then tobacco leaves were used for transient expression. After 2 days of dark treatment, 1 mmol of fluorescein (Promega, USA) was applied to the injection site of tobacco leaves and dark treated for six minutes. The signals were then captured with a CCD imaging instrument (Alliance, UK). Three independent leaves were observed. The method was described by Liu et al. (2021).

Arabidopsis transformation

The coding sequence of *EjWUSa* was cloned into the pFGC5941 vector with the special primers in Table S3. The constructed fusion vector was transformed into *Agrobacterium strain* GV3101, and then wild-type *Arabidopsis* (WT) were used for stable expression assay. The seeds of *35S::EjWUSa* transgenic *Arabidopsis* and WT were planted in the soil after treatment at 4°C for 24 h. Basta was used to screen *35S::EjWUSa* transgenic *Arabidopsis*. DNA was extracted from *35S::EjWUSa* transgenic *Arabidopsis* and WT leaves and then PCR was performed. RNA was extracted from *35S::EjWUSa* transgenic *Arabidopsis* and WT and then expression analysis was performed using special primers in Table S3. Finally, we obtained the T3 homozygous *35S::EjWUSa* transgenic lines. We counted the bolting time and the flowering time of *35S::EjWUSa* transgenic *Arabidopsis* and WT, respectively. Meanwhile, we counted the number of rosette leaves of *35S::EjWUSa* transgenic *Arabidopsis* and WT when the flowering shoot was 1 cm, respectively. All data were analyzed for significance by SPSS 26.0 software with One-way ANOVAs analysis.

Results

Identification and phylogenetic tree of *EjWOX* genes in loquat

The HMM of *HOX* superfamily was used as a query to search for candidate *EjHOX* genes in the loquat genome. A total of 134 candidate *EjHOX* genes were identified in the loquat genome, and the conserved domain of candidate *EjHOX* genes was analyzed in NCBI. The result showed 18 *EjWOX* genes were finally identified in the loquat (Figure S1). Phylogenetic analysis showed that the WOX proteins in loquat, apple, *Arabidopsis*, and tomato were similar (Figure 1). The *EjWOX* genes were classified into three well-supported clades (Figure 1). The WUS clade had the largest number of members, 11 in total, and the intermediate clade and ancient clade contained 4 and 3 members, respectively (Figure 1). The *EjWOX* genes were divided into nine subgroups based on phylogenetic analysis (Figure 1). Multiple sequence alignment showed that 13 amino acid residues are strictly conserved in the homeodomain of the EjWOX proteins, including Q and L in helix1, G in loop, P and L in helix2, G in turn, and N, V, W, F, Q, N, and R in helix3 (Figure 2A). An extra amino acid residue in the black box was observed in the homeodomain of the EjWOX proteins within the WUS subgroup (Figure 2A). The amino acid residue might be essential for their biological function. The WUS-box was

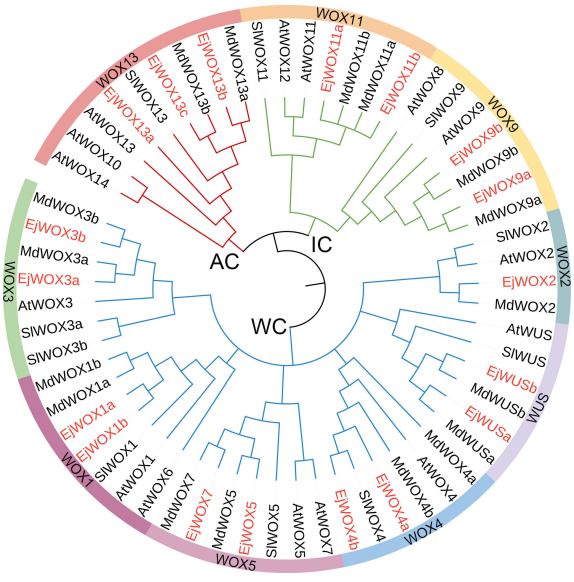


FIGURE 1
The phylogenetic tree including WOX proteins from *Eriobotrya japonica, Malus domestica, Solanum lycopersicum*, and *Arabidopsis thaliana*. Nine subgroups of the *WOX* family are represented by different colors. WC, WUS clade. IC, intermediate clade. AC, ancient clade. The WOX proteins in loquat are marked with red.

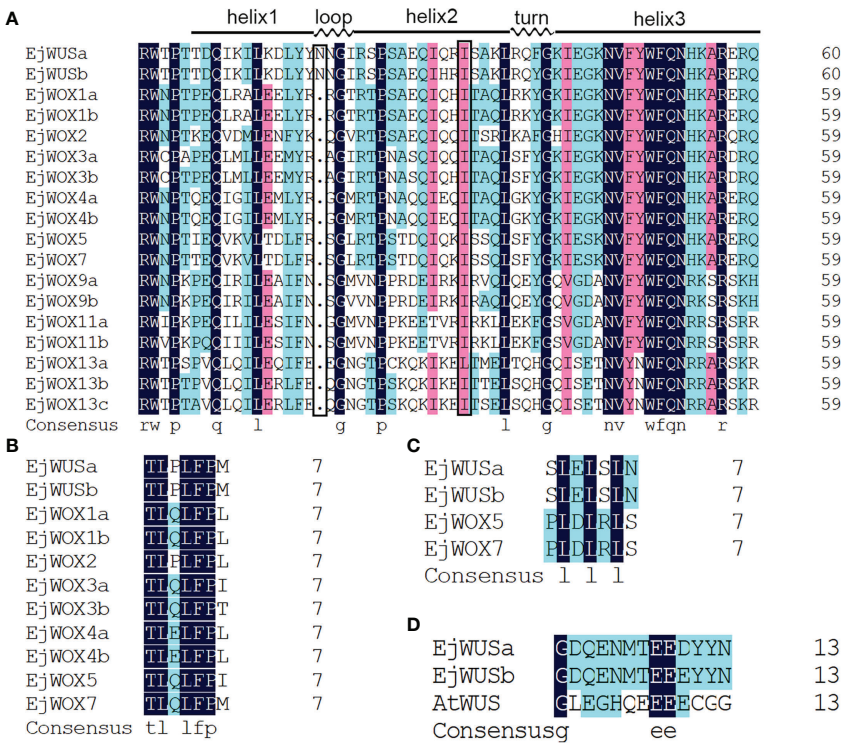


FIGURE 2
Multiple sequence alignment of WOX proteins in loquat. **(A)** Multiple sequence alignment of the homeodomain. An extra amino acid residue in the black box was observed in the homeodomain of the EjWOX proteins within the WUS subgroup. **(B)** Multiple sequence alignment of the WUS-box. **(C)** Multiple sequence alignment of the EAR-like motif. **(D)** Multiple sequence alignment of the acidic region. The identical residues are in black boxes.

found only in EjWOX proteins within the WUS clade (Figure 2B). The EAR-like motif was found only in EjWOX proteins within the WUS subgroup and WOX5 subgroup (Figure 2C). The acidic region was found only in EjWOX proteins within the WUS subgroup (Figure 2D).

Chromosomal localization, synteny analysis, and characterizations of the *EjWOX* genes in loquat

We analyzed the localization of 18 *EjWOX* genes on 17 loquat chromosomes. The 18 *EjWOX* genes were unevenly distributed on 12 of the 17 chromosomes (Figure 3A). No *EjWOX* genes were distributed on chromosomes 7, 8, 11, 12, and 13 (Figure 3A). Chromosomes 1, 2, 4, 10, 15, 16, and 17 had only one *EjWOX* gene (Figure 3A). Chromosomes 5, 6, 9, and 14 had two *EjWOX* genes (Figure 3A). Moreover, chromosome 3 had the highest number of *EjWOX* genes, with three *EjWOX* genes (Figure 3A). A total of 8 pairs of segmental duplication genes and 0 pairs of tandem duplication genes were identified in the loquat *WOX* family (Figure 3B). The length of EjWOX proteins ranged from 176 aa (EjWOX7) to 409 aa (EjWOX9a) with an average of 286.94 aa (Table 1). Their molecular weights ranged from 20.16 KDa (EjWOX7) to 45.14 KDa (EjWOX9a), with an average of 32.25 KDa (Table 1). In addition, these proteins might be mainly composed of basic amino acids with isoelectric points ranging from 4.71 (EjWOX13a) to 9.26 (EjWOX4b), with an average of 7.48 (Table 1).

Conserved motif and gene structure analysis of the *EjWOX* genes in loquat

The conserved motifs of the loquat WOX proteins were analyzed using MEME program and five different motifs were obtained (Figure 4A). All EjWOX proteins contain motif 1 and motif 2 (Figure 4A). Sequence analysis showed that motif 1 and motif 2 make up the homeodomain. All EjWOX proteins within the WUS clade contain motif 4 (Figure 4A). Sequence analysis showed that motif 4 is the WUS-box. All EjWOX proteins within the ancient clade contain motif 3 (Figure 4A). Except for EjWOX11b, all EjWOX proteins within the intermediate clade contain motif 5 (Figure 4A). EjWOX proteins within the same

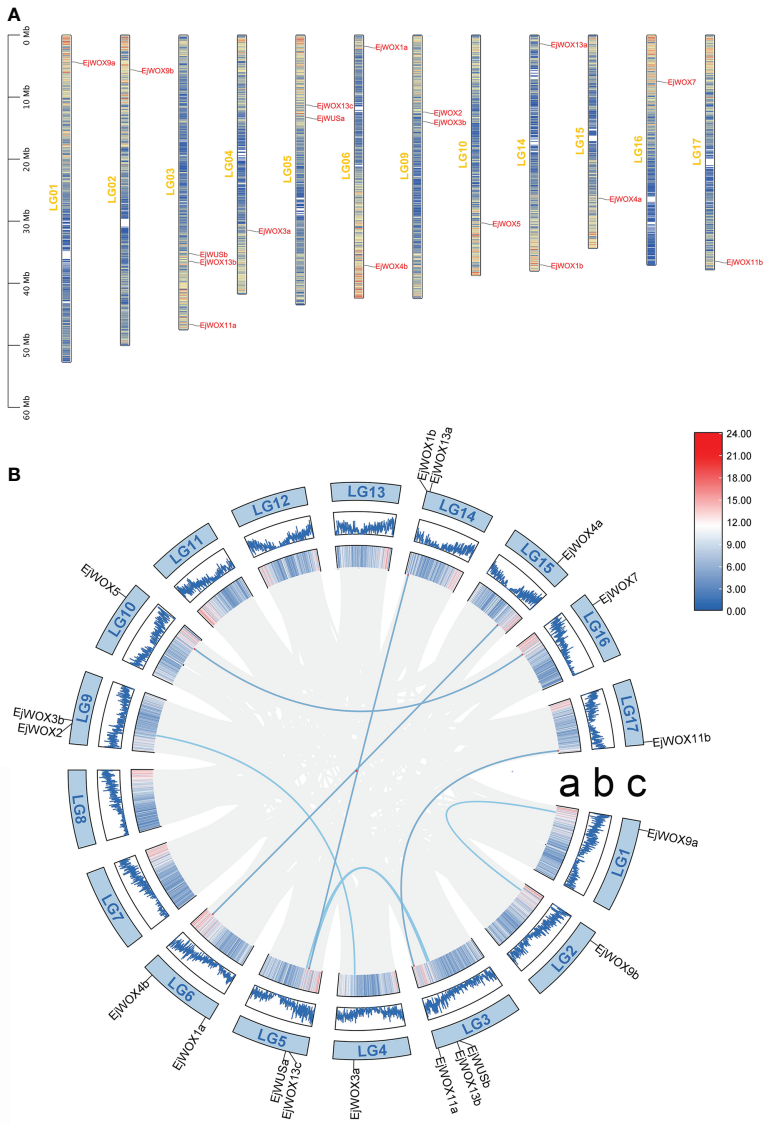


FIGURE 3
Chromosomal localization and synteny analysis of the *EjWOX* genes. **(A)** The chromosomal localization of the *EjWOX* genes. The chromosome numbers are marked on the left of each chromosome. Gene names are marked to the right of each chromosome. The black line with the scale on the left represents the length of the chromosomes. The colored lines within these columns represent gene density. **(B)** The synteny analysis of the *EjWOX* genes. Syntenic gene pairs are connected by blue lines. a: The colored lines indicate gene density. b: The width of the blue lines represents the gene density. c: LG01-LG17 represent 17 loquat chromosomes.

clade contained similar motifs, indicating that they might undertake similar biological functions (Figure 4A). We further analyzed the exon-intron structure of the *EjWOX* genes. All *EjWOX* genes contain introns (Figure 4B). The intron number of *EjWOX* genes varied from one to three (Figure 4B). Among the 18 *EjWOX* genes, there were 11 genes containing untranslated regions (Figure 4B). Only 7 *EjWOX* genes did not contain untranslated regions, i.e., *EjWUSa*, *EjWUSb*, *EjWOX2*, *EjWOX5*, *EjWOX7*, *EjWOX11a*, and *EjWOX11b* (Figure 4B). The number and length of exons, introns, and untranslated regions (UTRs) were conserved in *EjWOX* genes within the same subgroup (Figure 4B).

Temporal expression patterns of *EjWUSa* and *EjWUSb* in loquat flower buds

From July to November 2021, the loquat flower buds at 9 stages were collected from the 12-year-old triploid loquat 'Wuhezaoyu'. The RNA degradation of loquat flower buds at

TABLE 1 The characteristics of *EjWOX* genes in loquat.

Gene name	Gene locus	Chr.no.	Strand direction	Location	Protein		
					Lenth(aa)	Mol.Wt.(KDa)	pl
EjWUSa	Eja05G013550.1	5	–	10997929-10999501	328	36.06	6.67
EjWUSb	Eja03G017360.1	3	+	34685761-34687314	324	35.56	6.86
EjWOX1a	Eja06G001390.1	6	+	2303751-2306980	381	43.06	6.77
EjWOX1b	Eja14G019830.1	14	–	50634-53908	376	42.649	7.76
EjWOX2	Eja09G010840.1	9	+	10468783-10470302	271	30.57	8.81
EjWOX3a	Eja04G017350.1	4	+	29535577-29538433	244	28.15	7.71
EjWOX3b	Eja09G011670.1	9	–	11986178-11989035	237	27.28	8.97
EjWOX4a	Eja15G014170.1	15	–	27685794-27687603	239	26.73	9.26
EjWOX4b	Eja06G024530.1	6	–	37418904-37420581	238	26.75	9.24
EjWOX5	Eja10G018800.1	10	–	29983652-29984397	179	20.41	8.76
EjWOX7	Eja16G008720.1	16	+	6688246-6688915	176	20.16	8.76
EjWOX9a	Eja01G005310.1	1	–	4323964-4327524	409	45.14	7.88
EjWOX9b	Eja02G005930.1	2	–	4892997-4895908	407	45	7.19
EjWOX11a	Eja03G028280.1	3	–	46141451-46144547	302	32.85	5.48
EjWOX11b	Eja17G022460.1	17	–	36473574 36474613	274	30.6	8.45
EjWOX13a	Eja14G000850.1	14	+	1721638- 1724233	235	27.12	4.71
EjWOX13b	Eja03G018500.1	3	+	35974797- 35978213	277	31.65	5.62
EjWOX13c	Eja05G011640.1	5	–	9250942- 9253815	268	30.82	5.72

"+" and "-" represent strand direction.

the petal fall was serious, so the loquat flower buds at the petal fall were not suitable for qRT-PCR (Figure 5A, B). Transcriptome analysis of loquat flowers showed that *WUS* might be a floral meristem identity gene (Jing et al., 2020). We analyzed the expression levels of *EjWUSa* and *EjWUSb* in loquat flowers at different stages. In the early stages of flower development, the expression level of *EjWUSa* was significantly higher than that of *EjWUSb* (Figure 5C), suggesting that *EjWUSa* might play an important role in the transition from the vegetative apex to reproductive apex. Therefore, we further cloned *EjWUSa* from loquat and then investigated its role in loquat flowering.

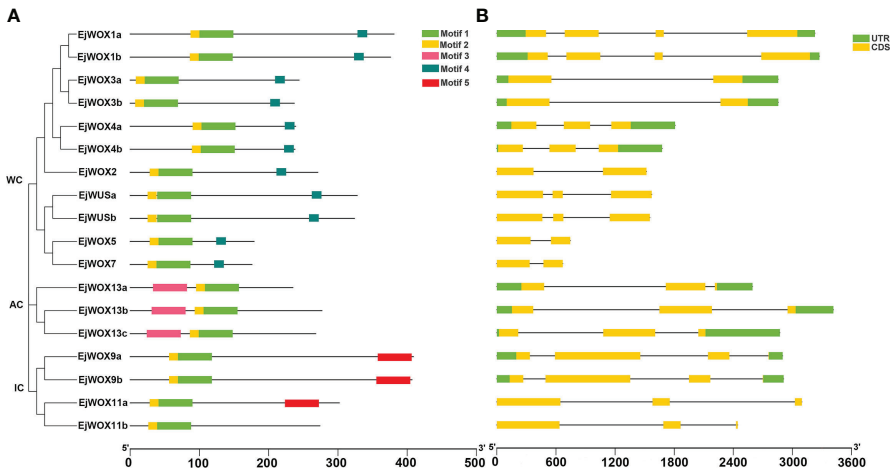


FIGURE 4
Conserved motif and gene structure analysis of loquat *WOX* genes. **(A)** Conserved motifs in EjWOX proteins are represented by colored boxes. **(B)** UTRs, exons, and introns are represented by green squares, yellow squares, and gray lines, respectively. Black lines indicate length. WC, WUS clade. IC, intermediate clade. AC, ancient clade.

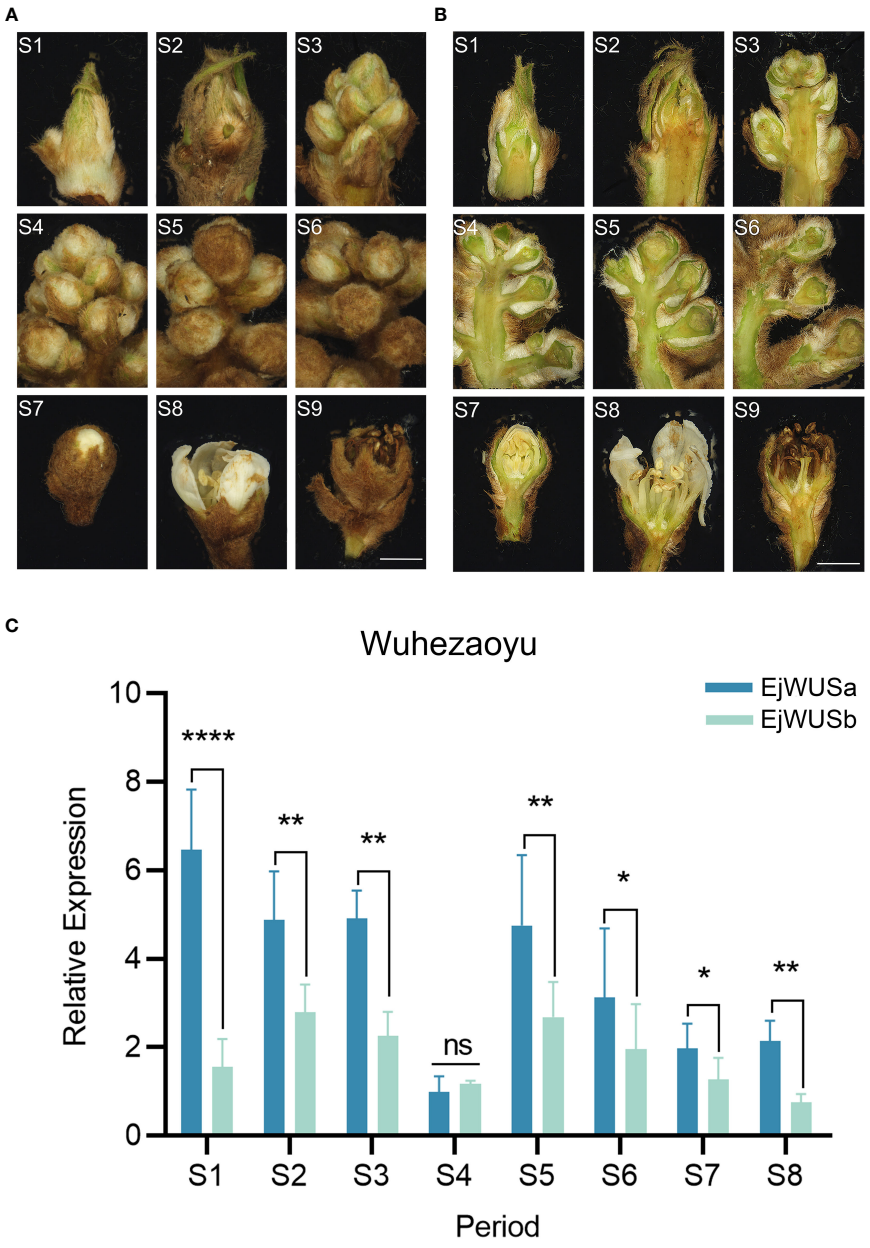


FIGURE 5
Inflorescence morphology and morphological anatomy of the triploid loquat 'Wuhezaoyu' and the temporal expression patterns of *EjWUSa* and *EjWUSb* in the triploid loquat 'Wuhezaoyu' flower buds. **(A, B)** Inflorescence morphology and morphological anatomy of the triploid loquat 'Wuhezaoyu'. S1: Vegetative apex; S2: Floral meristem initiation and flower bud differentiation; S3: Rapid differentiation of floral buds; S4: Panicle elongation; S5: Floral bud elongation with visible floral buds; S6: Elongation of branches in a panicle; S7: White corollas of floral buds; S8: Floral anthesis and full bloom; S9: Petal fall. Bars = 2000μm. **(C)** The temporal expression patterns of *EjWUSa* and *EjWUSb* in the triploid loquat 'Wuhezaoyu' flower buds. *Ejactin* as an internal control. Error bars indicate Standard Error (SE) from three biological replicates. Asterisks or n.s. indicate significant differences between *EjWUSa* and *EjWUSb*, *p<0.05, **p<0,01, **** P < 0.0001, ns, no significant difference, by One-way ANOVAs.

Isolation of *EjWUSa* from loquat

The *EjWUSa* was isolated from the cDNA of triploid loquat 'Wuhezaoyu' flowers (Figure S3), which contains three exons and two introns, but no untranslated region (Figure 4B). The coding sequence of *EjWUSa* is 987 bp, encoding 328 amino acids (Table 1). The EjWUSa contains four conserved domains from the N-terminal to the C-terminal, followed by the

homeodomain, the acidic region, the WUS-box, and the EAR-like motif (Figure 2). The *EjWUSa* promoter was cloned from the DNA of the triploid loquat 'Wuhezaoyu' leaves (Figure S3; Table S4). We further analyzed the binding elements on the *EjWUSa* promoter (Figure S4). The result showed that the *EjWUSa* promoter contains one CAT-box related to meristem expression (Figure S4) (Lin et al., 2022), suggesting that the *EjWUSa* might be involved in forming the shoot apical meristem. In addition, the *EjWUSa* promoter also contains two Box4s and two TCCC-motifs (Figure S4) (Chen and Qiu, 2020), indicating that the expression of *EjWUSa* might be regulated by light signals.

Subcellular localization of EjWUSa

The coding sequence (without stop codon) of *EjWUSa* was cloned into the pCAMBIA1300 vector to generate a fusion protein in tobacco cell. The green fluorescent protein (GFP) in the control group was localized in the nucleus and cell membrane, while the EjWUSa-GFP fusion protein in the experimental group was localized in the nucleus (Figure 6). It showed that the EjWUSa protein was localized in the nucleus, consistent with the characteristics of transcription factors.

The interactions between EjWUSa with EjWUSa and EjSTM

To investigate whether the protein interactions between WUS and WUS and STM are conserved across species, we isolated the *EjSTM* from the cDNA of the triploid loquat 'Wuhezaoyu' flower buds (Figure S3; Table S5). The BiFC assay and LCI assay were used to verify whether EjWUSa can form a dimer with EjSTM or EjWUSa in tobacco cells. In the experimental groups, strong signals were observed in tobacco cells, while no signals were observed in the control groups (Figure 7). This suggested that EjWUSa formed dimers with EjWUSa or EjSTM, respectively, which provided the basis for the formation of a complete YFP or LUC.

Arabidopsis transformation

To investigate the function of *EjWUSa*, an overexpression vector containing the coding sequence of the *EjWUSa* gene was constructed and transformed into wild-type *Arabidopsis* (WT). After Basta screening and PCR identification, we obtained T3 homozygous *35S::EjWUSa* transgenic lines (Figure S5). The expression level of *EjWUSa* in *35S::EjWUSa* transgenic lines was higher than in WT (Figure 8D). Under the same growing conditions, WT had about 12 rosette leaves, whereas the *35S::EjWUSa* transgenic lines had only seven to nine rosette leaves (Figures 8B, C). All *35S::EjWUSa* transgenic lines had approximately 10 days early bolting and 9 days early flowering compared to WT (Figures 8A, B). In conclusion, all *35S::EjWUSa* transgenic lines showed an early flowering phenotype compared to WT. Compared to WT, the *35S::EjWUSa* transgenic lines had no significant difference in morphological characteristics such as flower organs, leaf shape, siliques, stems, and leaves. From the above results, *EjWUSa* has the function of promoting *Arabidopsis* flowering.

Discussion

As a category of transcription factors regulating stem cell fate, WOX proteins are involved in many physiological processes

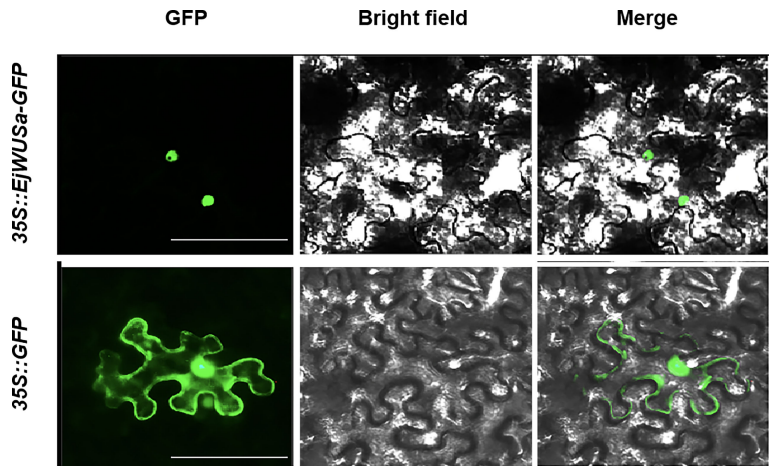


FIGURE 6
Subcellular localization of EjWUSa in tobacco leaves. GFP, GFP fluorescence channel; Merged, merged image of GFP and bright field. Bars = 50 μm.

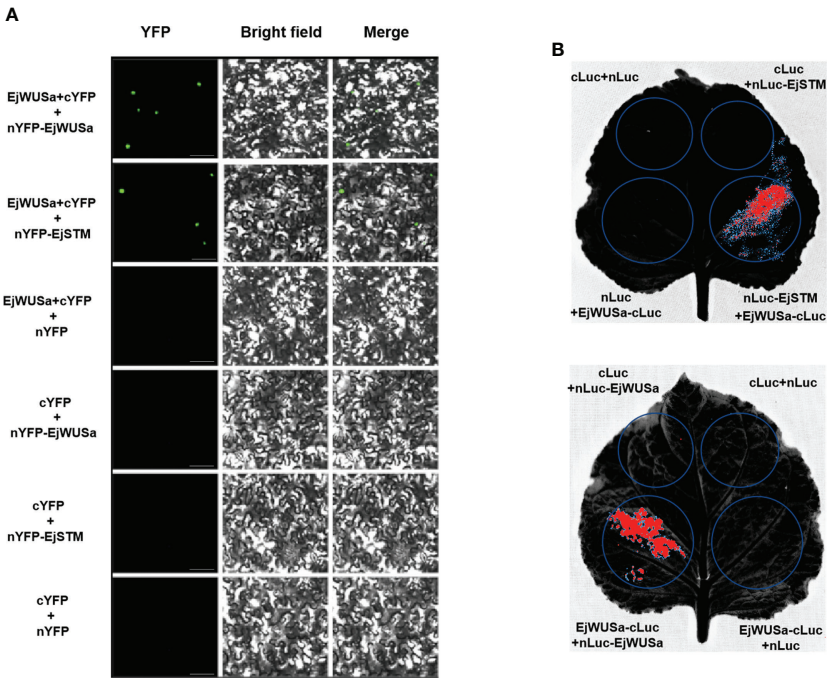


FIGURE 7
Protein interactions between EjWUSa with EjWUSa and EjSTM. **(A)** Protein interactions between EjWUSa with EjWUSa and EjSTM were demonstrated in the BIFC assay. The scale bar is 20 μm. **(B)** Protein interactions between EjWUSa with EjWUSa and EjSTM were demonstrated in the LCI assay.

related to plant growth and development (van der Graaff et al., 2009; Gambino et al., 2011; Lin et al., 2012; Kanchan and Sembi, 2020). Due to advances in omics technologies, genome-wide identification of the *WOX* gene family has been accomplished in several species, including *Arabidopsis*, sorghum, maize, tobacco, potato, walnut, and sweet orange. These plants contain 15, 11, 21, 10, 8, 14, and 8 *WOX* genes, respectively (2020; Vandenbussche et al., 2009; Zhang et al., 2010; Li et al., 2018; Shafique Khan et al., 2021). However, genome-wide identification and function of the *WOX* genes have not been reported in loquat. In the present study, we performed the genome-wide identification of the *WOX* gene family in loquat. Meanwhile, we further investigated the role of *EjWUSa* in loquat flowering.

A total of 18 *EjWOX* genes were identified in the loquat genome (Table S1). Like the *AtWOX* genes, the *EjWOX* genes can be divided into three clades or nine subgroups. Nevertheless, the homologs of *Arabidopsis AtWOX6*, *AtWOX10*, and *AtWOX14* could not be found in loquat (Figure 1). The substitutability of *WOX* genes might make their loss possible (Feng et al., 2021). Gene duplication is the main reason for gene family expansion (Feng et al., 2021). In our study, 8 pairs of segmental duplication genes and 0 pairs of tandem duplication genes were identified in the loquat *WOX* family (Figure 3B), suggesting that segmental duplications might be the main reason for the expansion of the loquat *WOX* family. The homeodomain is conserved in the *WOX* gene family among different species and maintains the functional integrity of *WOX* genes (Zhang et al., 2010; Kanchan and Sembi, 2020). As shown in Figure 2A, the 11th amino acid residue (In the black box) in helix 2 of the homeodomain is not strictly conserved in loquat WOX proteins, which differs from that of other species, such as apple (Figure S6), walnut, tomato, potato, rice, sorghum, maize, *Arabidopsis*, and poplar (Zhang et al., 2010; Li et al., 2018; Chang et al., 2019). We found a fragment of the EjWUS proteins consisting of 10 amino acids residues, rich in glutamate residues and aspartate residues, and with the same number of acidic residues as the AtWUS protein (Figure 2D). It should be considered as an acidic region. Therefore, the acidic region is not a conserved domain specific to the AtWUS protein. However, the acidic region was not detectable in WUS proteins in rice, sorghum, or maize (Zhang et al., 2010). The *EjWOX* genes within the ancient clade were more conserved in the number and length of exons, introns, and UTRs compared to those within the intermediate and modern clades, consistent with previous studies (Figure 4B) (Deveaux et al., 2008).

EjWUSa was shown to interact with EjWUSa and EjSTM (Figure 7), consistent with studies in *Arabidopsis* (Perales et al., 2016; Zhou et al., 2018). This suggested that protein interactions between WUS with WUS and STM might be conserved across

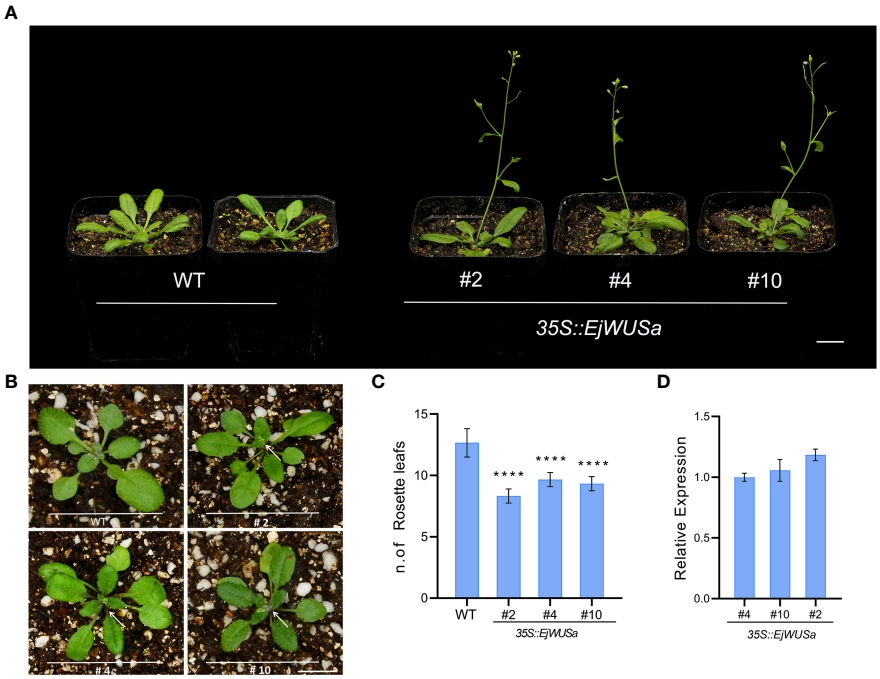


FIGURE 8
Overexpression of *EjWUSa* promotes *Arabidopsis* flowering. **(A)** *35S::EjWUSa* transgenic lines showed early flowering compared to WT. Bars = 1cm. **(B)** *35S::EjWUSa* transgenic lines showed early bolting compared to WT. Bars = 1cm. **(C)** The number of rosette leaves of *35S::EjWUSa* transgenic lines and WT while blooming. **(D)** The expression level of *EjWUSa* in *35S::EjWUSa* transgenic lines and WT. The leaves of *35S::EjWUSa* transgenic *Arabidopsis* and WT were collected after flowering and used for qRT-PCR analysis. Error bars indicate SE from three biological replicates. Asterisks indicate significant differences between *35S::EjWUSa* transgenic lines and WT, ****p < 0.0001, by One-way ANOVAs.

species. In *Arabidopsis*, AtWOX proteins were essential for embryonic patterning, stem cell maintenance, and organ formation (Ji et al., 2010). *AtWOX2* was involved in the formation of apical-basal axis (2008; Palovaara, 2010). The *wox3/prs1* mutant *Arabidopsis* exhibited short stature, abnormal sepal number and morphology, and the absence of stamens and stipules (Shimizu et al., 2009). *OsWOX3A*, which is the rice homolog of *AtWOX3*, was involved in the pleiotropic effects of organ development (Cho et al., 2013). As a homolog of *AtWOX9, Cymose Petunia WOX9* played an important role in the inflorescence development (Rebocho et al., 2008). *AtWOX13* and *AtWOX14* played important roles in the development of roots and the formation of floral organs (Deveaux et al., 2008). *AtWOX5* maintained the maintenance of stem cell homeostasis in the root apical meristem (RAM) (Sarkar et al., 2007; Nardmann et al., 2009; Tian et al., 2014). In *Arabidopsis* shoot and floral meristems, the *AtWUS* gene was required for stem cell identity (Schoof et al., 2000; Nardmann et al., 2007). When the *AtWUS* gene was mutated, it caused an early termination of inflorescence meristem (Laux et al., 1996). *Arabidopsis* overexpressing the *AtWUS* gene showed ectopic flower bud growth and callus-like formation (Xu et al., 2005). Functional verification of the *WUS* gene in other plants has been reported. For example, overexpression of the *GhWUS* gene increased the embryogenic callus formation in *Gossypium hirsutum* (Xiao et al., 2018). WUS protein has anti-virus activity by repressing MTase expression (Xu, 2021). In this study, *35S::EjWUSa* transgenic *Arabidopsis* exhibits early flowering (Figure 8). Unlike studies in *Arabidopsis*, we did not find the callus-like tissue on the stems of transgenic lines. Multiple sequence alignment showed that the similarity between EjWUSa and AtWUS was only 32.94%, which might lead to their functional differences (Figure S7). This also reflects the difference in genetic information between woody plants and herbs.

Conclusions

We reported the first genome-wide identification of the *WOX* family in perennial loquat. In the present study, 18 *EjWOX* genes were identified in the loquat genome. We further analyzed the evolutionary features and basic characteristics of the loquat *WOX* genes. A total of 8 pairs of segmental duplication genes and 0 pairs of tandem duplication genes were identified in the loquat *WOX* family, suggesting that segmental duplications might be the main reason for the expansion of the loquat *WOX* family. Compared with WT, *35S::EjWUSa* transgenic *Arabidopsis* exhibits early flowering.

Our study provides an important basis for further research on the function of *EjWOX* genes in the future and also facilitates the molecular breeding of loquat early-flowering varieties.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://ngdc.cncb.ac.cn/gwh, GWHBOTF00000000.

Author contributions

YY and MY performed the experiments and drafted the manuscript. XL contributed to morphology analysis of diploid loquat. YX, RH, and QX contributed to the data analysis. QG, DJ, and YX provided plant tissues, laboratory facilities, and project supervision. All authors approved the final draft of the manuscript.

Funding

This work was financially supported by The National Natural Science Foundation of China (Grant Numbers: 32102321), the National Key R&D Program of China (No. 2019YFD1000900), the Chongqing Science and Technology Commission (cstc2021jscx-gksbX0010 and cstc2021jcyj-msxmX1156), the Innovation Research Group Funds for Chongqing Universities (CXQT19005), the Chongqing Forestry Administration (YuLinKeYan2022-14), and the Characteristic Fruit Industry and Technology System Innovation Team of Chongqing Agriculture and Rural Affairs Commission: No.2020[3]01.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1024515/full#supplementary-material

References

Alvarez, J. M., Bueno, N., Cañas, R. A., Avila, C., Cánovas, F. M., and Ordás, R. J. (2018). Analysis of the WUSCHEL-RELATED HOMEOBOX gene family in *Pinus pinaster*: New insights into the gene family evolution. *Plant Physiol. Biochem.* 123, 304–318. doi: 10.1016/j.plaphy.2017.12.031

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37 (Web Server issue), W202–W208. doi: 10.1093/nar/gkp335

Bergonzi, S., Albani, M. C., Ver Loren van Themaat, E., Nordström, K. J. V., Wang, R., Schneeberger, K., et al. (2013). Mechanisms of age-dependent response to winter temperature in perennial flowering of *Arabis alpina*. *Sci. (New York N.Y.)* 340 (6136), 1094–1097. doi: 10.1126/science.1234116

Cao, S., Yang, Z., and Zheng, Y. (2013). Sugar metabolism in relation to chilling tolerance of loquat fruit. *Food Chem.* 136 (1), 139–143. doi: 10.1016/j.foodchem.2012.07.113

Chang, Y., Song, X., Zhang, Q., Liu, H., Bai, Y., Lei, X., et al. (2019). Genome-wide identification of WOX gene family and expression analysis during rejuvenational rhizogenesis in walnut (*Juglans regia* l.). *Forests* 11 (1), 16. doi: 10.3390/f11010016

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13 (8), 1194–1202. doi: 10.1016/j.molp.2020.06.009

Chen, S., and Qiu, G. (2020). Cloning and activity analysis of the promoter of nucleotide exchange factor gene ZjFes1 from the seagrasses *Zostera japonica*. *Sci. Rep.* 10 (1), 17291. doi: 10.1038/s41598-020-74381-6

Chen, W., Wang, P., Wang, D., Shi, M., Xia, Y., He, Q., et al. (2020). EjFRI, FRIGIDA (FRI) ortholog from *Eriobotrya japonica*, delays flowering in *Arabidopsis*. *Int. J. Mol. Sci.* 21 (3), E1087. doi: 10.3390/ijms21031087

Cho, S.-H., Yoo, S.-C., Zhang, H., Pandeya, D., Koh, H.-J., Hwang, J.-Y., et al. (2013). The rice narrow leaf2 and narrow leaf3 loci encode WUSCHEL-related homeobox 3A (OsWOX3A) and function in leaf, spikelet, tiller and lateral root development. *New Phytol.* 198 (4), 1071–1084. doi: 10.1111/nph.12231

Daum, G., Medzihradszky, A., Suzaki, T., and Lohmann, J. U. (2014). A mechanistic framework for noncell autonomous stem cell induction in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 111 (40), 14619–14624. doi: 10.1073/pnas.1406446111

Deveaux, Y., Toffano-Nioche, C., Claisse, G., Thareau, V., Morin, H., Laufs, P., et al. (2008). Genes of the most conserved WOX clade in plants affect root and flower development in *Arabidopsis*. *BMC Evol. Biol.* 8, 291. doi: 10.1186/1471-2148-8-291

Esumi, T., Tao, R., and Yonemori, K. (2005). Isolation of LEAFY and TERMINAL FLOWER 1 homologues from six fruit tree species in the subfamily maloideae of the rosaceae. *Sex. Plant Reprod.* 17 (6), 277–287. doi: 10.1007/s00497-004-0239-3

Feng, C., Zou, S., Gao, P., and Wang, Z. (2021). In silico identification, characterization expression profile of WUSCHEL-related homeobox (WOX) gene family in two species of kiwifruit. *PeerJ* 9, e12348. doi: 10.7717/peerj.12348

Gambino, G., Minuto, M., Boccacci, P., Perrone, I., Vallania, R., and Gribaudo, I. (2011). Characterization of expression dynamics of WOX homeodomain

transcription factors during somatic embryogenesis in *Vitis vinifera*. *J. Exp. Bot.* 62 (3), 1089–1101. doi: 10.1093/jxb/erq349

Ikeda, M., Mitsuda, N., and Ohme-Takagi, M. (2009). *Arabidopsis* WUSCHEL is a bifunctional transcription factor that acts as a repressor in stem cell regulation and as an activator in floral patterning. *Plant Cell* 21 (11), 3493–3505. doi: 10.1105/tpc.109.069997

Jiang, Y., Peng, J., Wang, M., Su, W., Gan, X., Jing, Y., et al. (2019a). The role of EjSPL3, EjSPL4, EjSPL5, and EjSPL9 in regulating flowering in loquat (*Eriobotrya japonica* lindl.). *Int. J. Mol. Sci.* 21 (1), E248. doi: 10.3390/ijms21010248

Jiang, Y., Peng, J., Zhang, Z., Lin, S., Lin, S., and Yang, X. (2019b). The role of EjSVPs in flower initiation in *Eriobotrya japonica*. *Int. J. Mol. Sci.* 20 (23), E5933. doi: 10.3390/ijms20235933

Jiang, Y., Peng, J., Zhu, Y., Su, W., Zhang, L., Jing, Y., et al. (2019c). The role of EjSOC1s in flower initiation in *Eriobotrya japonica*. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00253

Jing, D., Chen, W., Hu, R., Zhang, Y., Xia, Y., Wang, S., et al. (2020). An integrative analysis of transcriptome, proteome and hormones reveals key differentially expressed genes and metabolic pathways involved in flower development in loquat. *Int. J. Mol. Sci.* 21 (14), E5107. doi: 10.3390/ijms21145107

Jing, D., Chen, W., Xia, Y., Shi, M., Wang, P., Wang, S., et al. (2020). Homeotic transformation from stamen to petal in *Eriobotrya japonica* is associated with hormone signal transduction and reduction of the transcriptional activity of EjAG. *Physiol. Plant.* 168 (4), 893–908. doi: 10.1111/ppl.13029

Ji, J., Shimizu, R., Sinha, N., and Scanlon, M. J. (2010). Analyses of WOX4 transgenics provide further evidence for the evolution of the WOX gene family during the regulation of diverse stem cell functions. *Plant Signa Behav.* 5 (7), 916–920. doi: 10.4161/psb.5.7.12104

Kanchan, M., and Sembi, J. (2020). Comparative transcriptomic analyses of four phalaenopsis species to identify and characterize the WUSCHEL -related homeobox (WOX) gene family. *Biotechnologia* 101 (4), 309–322. doi: 10.5114/bta.2020.100423

Laux, T., Mayer, K. F., Berger, J., and Jürgens, G. (1996). The WUSCHEL gene is required for shoot and floral meristem integrity in *Arabidopsis*. *Dev. (Cambridge England)* 122 (1), 87–96. doi: 10.1242/dev.122.1.87

Li, X., Hamyat, M., Liu, C., Ahmad, S., Gao, X., Guo, C., et al. (2018). Identification and characterization of the WOX family genes in five solanaceae species reveal their conserved roles in peptide signaling. *Genes* 9 (5), E260. doi: 10.3390/genes9050260

Li, Z., Liu, D., Xia, Y., Li, Z., Jing, D., Du, J., et al. (2020). Identification of the WUSCHEL-related homeobox (WOX) gene family, and interaction and functional analysis of TaWOX9 and TaWUS in wheat. *Int. J. Mol. Sci.* 21 (5), E1581. doi: 10.3390/ijms21051581

Lin, H., Niu, L., McHale, N. A., Ohme-Takagi, M., Mysore, K. S., and Tadege, M. (2012). Evolutionarily conserved repressive activity of WOX proteins mediates leaf blade outgrowth and floral organ development in plants. *Proc. Natl. Acad. Sci. U. S. A.* 110 (1), 366–371. doi: 10.1073/pnas.1215376110

Lin, M., Yan, J., Ali, M. M., Wang, S., Tian, S., Chen, F., et al. (2022). Isolation and functional characterization of a green-tissue promoter in japonica rice (*Oryza sativa* subsp. *Japonica*). *Biology* 11 (8), 1092. doi: 10.3390/biology11081092

Liu, Y., Song, H., Liu, Z., Hu, G., and Lin, S. (2013). Molecular characterization of loquat EjAP1 gene in relation to flowering. *Plant Growth Regul.* 70 (3), 287–296. doi: 10.1007/s10725-013-9800-0

Liu, Y., Wen, L., Shi, Y., Su, D., Lu, W., Cheng, Y., et al. (2021). Stress-responsive tomato gene SlGRAS4 function in drought stress and abscisic acid signaling. *Plant Sci.* 304, 110804. doi: 10.1016/j.plantsci.2020.110804

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 (San Diego Calif.)* 25 (4), 402–408. doi: 10.1006/meth.2001.1262

Lopes, F. L., Galvan-Ampudia, C., and Landrein, B. (2021). WUSCHEL in the shoot apical meristem: old player, new tricks. *J. Exp. Bot.* 72 (5), 1527–1535. doi: 10.1093/jxb/eraa572

Munir, S., Khan, M. R. G., Song, J., Munir, S., Zhang, Y., Ye, Z., et al. (2016). Genome-wide identification, characterization and expression analysis of calmodulin-like (CML) proteins in tomato (*Solanum lycopersicum*). *Plant Physiol. Biochem.* 102, 167–179. doi: 10.1016/j.plaphy.2016.02.020

Nardmann, J., Reisewitz, P., and Werr, W. (2009). Discrete shoot and root stem cell-promoting WUS/WOX5 functions are an evolutionary innovation of angiosperms. *Mol. Biol. Evol.* 26 (8), 1745–1755. doi: 10.1093/molbev/msp084

Nardmann, J., Zimmermann, R., Durantini, D., Kranz, E., and Werr, W. (2007). WOX gene phylogeny in poaceae: A comparative approach addressing leaf and embryo development. *Mol. Biol. Evol.* 24 (11), 2474–2484. doi: 10.1093/molbev/msm182

Palovaara, J. (2008). Conifer WOX-related homeodomain transcription factors, developmental consideration and expression dynamic of WOX2 during *Picea abies* somatic embryogenesis. *Plant Mol. Biol.* 66 (5), 533–549. doi: 10.1007/s11103-008-9289-5

Palovaara, J. (2010). Comparative expression pattern analysis of WUSCHEL-related homeobox 2 (WOX2) and WOX8/9 in developing seeds and somatic embryos of the gymnosperm *Picea abies*. *New Phytol.* 188 (1), 122–135. doi: 10.1111/j.1469-8137.2010.03336.x

Perales, M., Rodriguez, K., Snipes, S., Yadav, R. K., Diaz-Mendoza, M., and Reddy, G. V. (2016). Threshold-dependent transcriptional discrimination underlies stem cell homeostasis. *Proc. Natl. Acad. Sci. U. S. A.* 113 (41), E6298–E6306. doi: 10.1073/pnas.1607669113

Rebocho, A. B., Bliek, M., Kusters, E., Castel, R., Procissi, A., Roobeek, I., et al. (2008). Role of EVERGREEN in the development of the cymose petunia inflorescence. *Dev. Cell* 15 (3), 437–447. doi: 10.1016/j.devcel.2008.08.007

Reig, C., Gil-Muñoz, F., Vera-Sirera, F., García-Lorca, A., Martínez-Fuentes, A., Mesejo, C., et al. (2017). Bud sprouting and floral induction and expression of FT in loquat [*Eriobotrya japonica* (Thunb.) lindl.]. *Planta* 246 (5), 915–925. doi: 10.1007/s00425-017-2740-6

Sarkar, A. K., Luijten, M., Miyashima, S., Lenhard, M., Hashimoto, T., Nakajima, K., et al. (2007). Conserved factors regulate signalling in *Arabidopsis thaliana* shoot and root stem cell organizers. *Nature* 446 (7137), 811–814. doi: 10.1038/nature05703

Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F., Jürgens, G., and Laux, T. (2000). The stem cell population of *Arabidopsis* shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* 100 (6), 635–644. doi: 10.1016/s0092-8674(00)80700-x

Shafique Khan, F., Zeng, R.-F., Gan, Z.-M., Zhang, J.-Z., and Hu, C.-G. (2021). Genome-wide identification and expression profiling of the WOX gene family in *Citrus sinensis* and functional analysis of a CsWUS member. *Int. J. Mol. Sci.* 22 (9), 4919. doi: 10.3390/ijms22094919

Shimizu, R., Ji, J., Kelsey, E., Ohtsu, K., Schnable, P. S., and Scanlon, M. J. (2009). Tissue specificity and evolution of meristematic WOX3 function. *Plant Physiol.* 149 (2), 841–850. doi: 10.1104/pp.108.130765

Su, Y. H., Zhou, C., Li, Y. J., Yu, Y., Tang, L. P., Zhang, W. J., et al. (2020). Integration of pluripotency pathways regulates stem cell maintenance in the *Arabidopsis* shoot meristem. *Proc. Natl. Acad. Sci. U. S. A.* 117 (36), 22561–22571. doi: 10.1073/pnas.2015248117

Tian, H., Wabnik, K., Niu, T., Li, H., Yu, Q., Pollmann, S., et al. (2014). WOX5-IAA17 feedback circuit-mediated cellular auxin response is crucial for the patterning of root stem cell niches in *Arabidopsis*. *Mol. Plant* 7 (2), 277–289. doi: 10.1093/mp/sst118

Tvorogova, V. E., Krasnoperova, E. Y., Potsenkovskaia, E. A., Kudriashov, A. A., Dodueva, I. E., and Lutova, L. A. (2021). [What does the WOX say? review of pegulators, targets, partners]. *Mol. Biol.* 55 (3), 362–391. doi: 10.31857/S0026898421030174

Vandenbussche, M., Horstman, A., Zethof, J., Koes, R., Rijpkema, A. S., and Gerats, T. (2009). Differential recruitment of WOX transcription factors for lateral development and organ fusion in petunia and *Arabidopsis*. *Plant Cell* 21 (8), 2269–2283. doi: 10.1105/tpc.109.065862

van der Graaff, E., Laux, T., and Rensing, S. A. (2009). The WUS homeobox-containing (WOX) protein family. *Genome Biol.* 10 (12), 248. doi: 10.1186/gb-2009-10-12-248

Wang, Y. (2021). A draft genome, resequencing, and metabolomes reveal the genetic background and molecular basis of the nutritional and medicinal properties of loquat (*Eriobotrya japonica* (Thunb.) lindl). *Hortic. Res.* 8 (1), 231. doi: 10.1038/s41438-021-00657-1

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., et al. (1999). Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol. (Clifton N.J.)* 112, 531–552. doi: 10.1385/1-59259-584-7:531

Xiao, Y., Chen, Y., Ding, Y., Wu, J., Wang, P., Yu, Y., et al. (2018). Effects of GhWUS from upland cotton (*Gossypium hirsutum* l.) on somatic embryogenesis and shoot regeneration. *Plant Sci.* 270, 157–165. doi: 10.1016/j.plantsci.2018.02.018

Xu, L. (2021). WUSCHEL: The versatile protein in the shoot apical meristem. *Sci. China Life Sci.* 64 (1), 177–178. doi: 10.1007/s11427-020-1870-4

Xu, Y.-Y., Wang, X.-M., Li, J., Li, J.-H., Wu, J.-S., Walker, J. C., et al. (2005). Activation of the WUS gene induces ectopic initiation of floral meristems on mature stem surface in *Arabidopsis thaliana*. *Plant Mol. Biol.* 57 (6), 773–784. doi: 10.1007/s11103-005-0952-9

Yadav, R. K. (2012). WUSCHEL protein movement and stem cell homeostasis. *Plant Signaling Behav.* 7 (5), 592–594. doi: 10.4161/psb.19793

Zhang, L., Jiang, Y., Zhu, Y., Su, W., Long, T., Huang, T., et al. (2019). Functional characterization of GI and CO homologs from *Eriobotrya deflexa* nakai forma *koshunensis*. *Plant Cell Rep.* 38 (5), 533–543. doi: 10.1007/s00299-019-02384-3

Zhang, L., Yu, H., Lin, S., and Gao, Y. (2016). Molecular characterization of FT and FD homologs from *Eriobotrya deflexa* nakai forma *koshunensis*. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00008

Zhang, X., Zong, J., Liu, J., Yin, J., and Zhang, D. (2010). Genome-wide analysis of WOX gene family in rice, sorghum, maize, *Arabidopsis* and poplar. *J. Integr. Plant Biol.* 52 (11), 1016–1026. doi: 10.1111/j.1744-7909.2010.00982.x

Zhou, Y., Yan, A., Han, H., Li, T., Geng, Y., Liu, X., et al. (2018). HAIRY MERISTEM with WUSCHEL confines CLAVATA3 expression to the outer apical meristem layers. *Sci. (New York N.Y.)* 361 (6401), 502–506. doi: 10.1126/science.aar8638

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*. *Sci. (New York N.Y.)* 340 (6136), 1097–1100. doi: 10.1126/science.1234340

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 February 2023
DOI 10.3389/fpls.2023.1093661

OPEN ACCESS

EDITED BY
Zhongxiong Lai,
Fujian Agriculture and Forestry University, Fuzhou, China

REVIEWED BY
Jianqing Chen,
Fujian Agriculture and Forestry University, China
Juyou Wu,
Nanjing Agricultural University, China

*CORRESPONDENCE
Chongchong Yan
527927758@qq.com
Yongjie Qi
anhuiqyj@163.com
Yongping Cai
chenyu19871215@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Development and EvoDevo, a section of the journal Frontiers in Plant Science

RECEIVED 22 November 2022
ACCEPTED 24 January 2023
PUBLISHED 08 February 2023

CITATION
Yan C, Zhang N, Xu C, Jin Q, Qi Y and Cai Y (2023) Effects on stone cell development and lignin deposition in pears by different pollinators.
Front. Plant Sci. 14:1093661.
doi: 10.3389/fpls.2023.1093661

Effects on stone cell development and lignin deposition in pears by different pollinators

Chongchong Yan[1,2*†], Nan Zhang[3,4†], Chao Xu[3], Qing Jin[3], Yongjie Qi[1,2*] and Yongping Cai[3*]

[1]Anhui Academy of Agricultural Sciences, Hefei, Anhui, China, [2]Insititute of Horticulture, Anhui Academy of Agricultural Sciences, Key Laboratory of Horticultural Crop Germplasm innovation and Utilization (Co-construction by Ministry and Province), Hefei, Anhui, China, [3]School of Life Science, Anhui Agricultural University, Hefei, Anhui, China, [4]College of Health and Elderly, Anhui Vocational College of City Management, Hefei, Anhui, China

Introduction: The pear pulp is formed by the development of the ovary wall, which is the somatic cell of the female parent, and its genetic traits are identical to those of the female parent, so that its phenotypic traits should also be identical to those of the female parent. However, the pulp quality of most pears, especially the stone cell clusters (SCCs) number and degree of polymerization (DP), were significantly affected by the paternal type. Stone cells are formed by the deposition of lignin in parenchymal cell (PC) walls. Studies on the effect of pollination on lignin deposition and stone cell formation in pear fruit have not been reported. Methods: In this study, 'Dangshan Su' (*P. bretschneideri* Rehd.) was selected as the mother tree, while 'Yali' (*P. bretschneideri* Rehd.) and 'Wonhwang' (*P. pyrifolia* Nakai.) were used as the father trees to perform cross-pollination. We investigated the effects of different parents on SCCs number and DP, and lignin deposition by microscopic and ultramicroscopic observation.

Results and Discussion: The results showed that the formation of SCCs proceeds was consistent in DY and DW, but the SCC number and DP in DY were higher than that in DW. Ultramicroscopy revealed that the lignification process of DY and DW were all from corner to rest regions of the compound middle lamella and the secondary wall, with lignin particles deposited along the cellulose microfibrils. They were alternatively arranged until they filled up the whole cell cavity to culminate in the formation of stone cells. However, the compactness of the wall layer of cell wall was significantly higher in DY than in DW. We also found that the pit of stone cell was predominantly single pit pair, they transported degraded material from the PCs that were beginning to lignify out of the cells. Stone cell formation and lignin deposition in pollinated pear fruit from different parents were consistent, but the DP of SCCs and the compactness of the wall layer were higher in DY than that in DW. Therefore, DY SCC had a higher ability to resist the expansion pressure of PC.

KEYWORDS

'Dangshan Su' pear, pollination, stone cell clusters, microscopy, ultramicroscopy

1 Introduction

Pyrus bretschneideri cv. 'Dangshan Su' is a Chinese pear species widely grown in China and throughout Asia (Xu, 2009). Most pears are self-incompatible fruits with obvious xenia, and the fruit quality is affected by the source of pollination. The effect of pollination source can be seen in attributes such as fruit hardness and sugar, acid, vitamin C, and SCC contents (Cai et al., 2010; Cheng et al., 2020; Mumtaz et al., 2020). We previously pollinated 'Dangshan Su' with pollen from two different varieties of pears and found that pollen from different parents significantly affected the expression of microRNAs and proteins related to lignin metabolism in pear fruit (Cheng et al., 2017; Li et al., 2018), thus affecting lignin deposition, which in turn affected stone cell formation in pear fruit and ultimately the taste of pear fruit.

Stone cells, also known as sclereids, are responsible for affecting pear quality and can be present in an individual fruit or in groups of fruit; groups of stone cells are called SCCs (Jin et al., 2013). Pear stone cells are classified as short stone cells, which are sclerenchyma cells with high levels of lignin and cellulose (Cai et al., 2010; Yan et al., 2014). These stone cells differentiate from PCs after the secondary wall of PCs deposit on their primary wall (Cheng et al., 2019). Stone cells develop through secondary thickening of the cell wall and lignin deposition. Uneven thickening of the secondary wall causes single pit to start to form(Lian et al., 2020; Chukhchin et al., 2021). Lignin deposition similar to numerous discrete granules first occur at the corner of the middle lamella (ML) within various primary cell walls. Biosynthesis, transport, and deposition of lignin are closely associated with stone cell development and influence the formation of stone cells in pears(Tao et al., 2009; Cai et al., 2010). Lignification extends from the corner to the remaining regions of the ML and the secondary wall simultaneously in the xylem cells of Populus tomentosa Carr (Guo et al., 2015). Due to the lack of relevant studies, it remains unclear how lignin deposits in the stone cell wall and what contributes to its secondary thickening during the development of pear fruit.

Many pollens can be used to fertilize the 'Dangshan Su' in Dangshan, Anhui Province, China, such as 'Wonhwang' (P. pyrifolia Nakai), 'Jingbaili' (P. ussuriensis Maxim.), 'bali' (P. communis Bartlett), and 'yali' (P. bretschneideri Rehd) (Cheng et al., 2017). Since the two varieties, 'Yali' and 'Wonhwang', are also widely grown in Dangshan. Therefore, the local area mainly use them as the parent to pollinate the 'Dangshan Su'. Thus, the pollen of 'Yali' and 'Wonhwang' was used to fertilize 'Dangshan Su' in this study. We compared the influence of these pollination types on the stone cell development and lignin deposition of 'Dangshan Su' fruits.

The results of this study can not only clarify the effects of pollination types on stone cell formation in pear, but also provide new insights into the mechanism of secondary wall formation of stone cells in pear. These results provide a theoretical foundation for selecting optimal pollination varieties for the 'Dangshan Su' pear, which could potentially improve the fruit quality.

Abbreviations: DY, Dangshan Su × Yali; DW, Dangshan Su × Wonhwang; SCC, stone cell cluster; DP, degree of polymerization; PC,parenchymal cell; DAF, Day after flowering; CML, compound middle lamella; CMfs, cellulose microfibrils; ML, middle lamella; SEM, Scanning electron microscopy; TEM, Transmission electron microscopy; CC, corner of cell; CIL, composite intercellular layer; CML, compound middle lamella; S1, S1 layer of the secondary wall; S1L, layer between S1 and S2; S2, S2 layer of the secondary wall; S2L, layer between S2 and S3; S3, S3, layer of the secondary wall.

2 Materials and methods

2.1 Materials

Fruits were obtained from pear trees grown on a farm (34°16'N 116°29'E) at the experimental site of the Institute of Horticulture, Anhui Academy of Agricultural Sciences, Dangshan County, Anhui Province, China. The 'Dangshan Su' female parent was fertilized with pollen from *P. bretschneideri.* 'Yali' (DY) and *P. pyrifolia.* 'Wonhwang' (DW). Buds on the short branches of 'Dangshan Su' with similar developmental stages and a similar size were selected from the mid-crown area on the south side of each tree. The stamens were removed from the buds and fertilized with the same amount of pollen from 'Yali' or 'Wonhwang'. We also selected pollen from buds with similar developmental stages and similar sizes. The newly pollinated buds were covered with yellow single layer bags for seven days after pollination. Previous reports demonstrated that pear stone cell development occurs from 15 DAF to 67 DAF (Cai et al., 2010). Therefore, equally sized pear fruits were sampled starting from 15 DAF. A total of 9 developmental periods were sampled: 15 DAF, 23 DAF, 31 DAF, 39 DAF, 47 DAF, 55 DAF, 63 DAF, 87 DAF and the mature period (145 DAF). Samples of pear fruit with similar size were collected. The fruits were frozen and brought back to the lab. A part of the pear samples were fixed with FAA and 3.5% formaldehyde and then stored under low temperature. Stone cell development was observed by microscopy and lignin deposition was observed by ultra-microscopy.

2.2 Observation of stone cell clusters

The morphology of the SCCs was observed according to the method of Yan et al.(Yan et al., 2014). Transverse sections of pear fruits were manually prepared and stained with 1.0% phloroglucinol and 1.0 M hydrochloric acid as described in the Wiesner lignin staining method (Escamez et al., 2017). The size and distribution of SCCs were observed after chromogenesis.

2.3 Microscopy of stone cell development and lignin deposition

The pulp, between 2 mm inside of the peel and 0.5 cm outside of the stone, was cut with a sharp blade to the appropriate size. The pulp cubes were subsequently fixed with FAA fixative (50 ml of 50% ethanol, 5 ml of acetic acid, and 5 ml of formalin) and manually sliced. One drop of 1 M Hcl was placed onto the sample on a slide, which was then stained with a drop of 0.1% phloroglucinol/ethanol. To further examine the lignin deposition in the stone cell wall during the development of the pear fruit, sections were treated with the Wiesner reagent (phloroglucinol - HCl staining) and observed under an oil immersion microscope (× 630) (Clifford, 1974; Tao

et al., 2009). A slide was microscopically examined, photographed and statistical stone cell mass size and distribution (Cai et al., 2010). The experiment was repeated six times, with three samples observed each time.

2.4 Scanning electron microscopy of stone cell distribution

Scanning electron microscopy (SEM) of stone cells was performed according to the method of Xu et al. (Xu et al., 2006). First, the fruit pulp sampled was screened from 2.0 mm under the peel to 0.5 mm outside the core was cut into small pieces (0.5 cm × 0.4 cm × 0.2 cm). The pulp was fixed for 1 h in 3.5% glutaraldehyde in phosphate buffer under low vacuum, followed by 2.5 hours at atmospheric pressure. Each sample was then fixed, washed and dried and 10 nm gold particles were plated onto the surface of the sample using a sputter coater. The samples were observed, photographed and content the stone cells number under the SEM (S-3000N, Hitachi, Japan) under high vacuum. The applied voltage between 6.00 kV 40 × and 6.00 kV 50 × and the working distance between 14.3 mm and 15.3 mm.

2.5 Ultramicroscopy of lignin deposition

Samples were stained with potassium permanganate ($KMnO_4$) and examined using transmission electron microscopy (TEM). The pulp from 2 mm inside of the peel to 0.5 cm outside of the stone was cut with a sharp blade to generate blocks of 0.5 cm × 0.2 cm × 0.2 cm. The blocks were immediately placed into fixation solution containing 3.5% glutaraldehyde and 0.2 M phosphate (pH 7.2), fixed under low vacuum for 1 h, followed by atmospheric pressure for 2.5 h. The samples were then washed with phosphate buffer and fixed in 1% osmium tetroxide (in 0.2 M phosphate buffer, pH 7.2) at 4°C overnight. The samples were subsequently dehydrated in a series of Spurr resin-water mixtures (1:2, 1:1 and 2:1), supplied by Marsys-Tech Co. Ltd., (Beijing, China, catalog number S2690), before being treated twice with pure Spurr resin. They were then embedded and aggregated, and cut into ultrathin sections with an LKB-2188 microtome. The ultrathin sections were collected with a 100 - mesh copper net. Some of the sections were double stained with uranyl acetate and lead citrate. The remaining sections were stained with 1% $KMnO_4$ solution (in 0.1% sodium citrate) for 3 min (Yin et al., 2003; Day et al., 2005). The stained sections were examined and photographed under a TEM (Philips EM-400T). The experiment was repeated six times, with one sample observed each time.

3 Results

3.1 Stone cell clusters distribution and number in DY and DW

After the Wiesner reaction, it could be revealed from the fruit sections that pear SSCs were stained rose red, in contrast to their surrounding PCs that were barely stained (Figures A1, B1). SSCs and PCs were arranged in a mosaic pattern such that SSCs were embedded in PCs (Shutian Tao et al., 2009). However, SSCs were not evenly distributed among PCs, they concentrated radially near the fruit center. The distribution of SCCs near the fruit center was also significantly more intensive than that in other areas. At the same time, we found that the nuclei of DY were smaller than DW, but the density and distribution range of SCCs were higher than that of DW. Observation of DY and DW using SEM, we can found that the unit

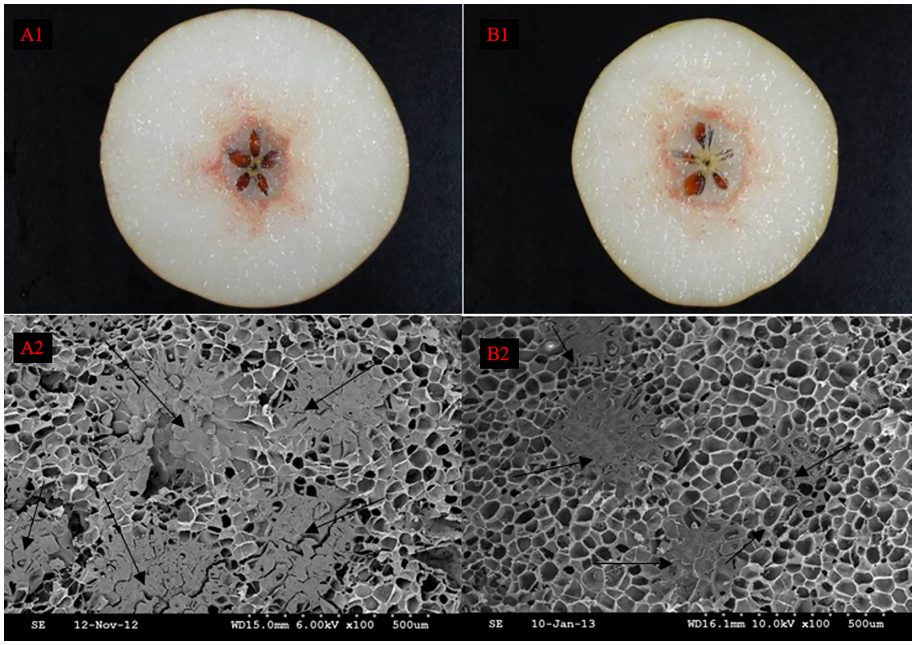


FIGURE 1
The distribution of stone cell cluster of DY and DW. **A1** Cross section staining of DY, **A2** Scanning electron microscopy of DY, **B1** Cross section staining of DW, **B2** Scanning electron microscopy of DW.

vision of DY showed that the number of the SCC was five, whereas in the DW it was only four, which also suggested that the distribution of SCCs was more intensive in DY than in DW (Figures A2, B2). This result is consistent with the results obtained from Wiesner reaction. It showed that pollination by different parents could change the distribution of SCCs in pear fruit.

3.2 Comparison the number and area of stone cell clusters in DY and DW

In both DY and DW, the area of SCCs initially increased and subsequently decreased over the course of fruit development (Figures 2, 3). At 23 DAF, the major SCCs in the fruits of DY and DW were in the primitive stage with loose stone cell DP. The SCC area of the DY and DW increased rapidly from 23 to 55 DAF. From 23 to 55 DAF, the SCC area of DY and DW increased rapidly, from 0.02 mm^2 to nearly 0.1 mm^2, and the formation of SCCs was largely complete at 55 DAF, after which it slowly decreased to about 0.06 mm^2. The number of SCCs was highest from 23 to 31 DAF and then gradually decreased. This might be due to the small size of pear fruit in the early stage of fruit development, and the number of SCCs in the unit field of view was more. With the development, the volume of fruit continues to increase, which leaded to the decrease of SCCs in the unit field of view. This was consistent with the findings of Lee et al. (Lee et al., 2006; Nie et al., 2009).

During pear fruit development, the number of SCCs in DW was higher than that in DY only at 31, 47 and 63 DAF. The area of DW SCCs was higher than that of DY only in 31, 39, 47 and 63 DAF (Figure 4). It could be seen that when the area of SCCs was higher in DY, the number of SCCs was also higher. The results indicated that the influence of pollen sensation on the number and area of SCCs in pear fruit was spatiotemporal.

3.3 Comparison of lignin deposition and stone cell development in DY and DW

Lignin deposition and transport were closely related to lignification of stone cells (Jin et al., 2013). The cell wall of pear fruit stone cells were generally divided into middle Lamella (ML), primary wall (PCW) and secondary wall (S). The secondary wall was divided into outer secondary wall (S1), middle secondary wall (S2) and inner secondary wall (S3). The composite middle lamella (CML) was formed between the middle lamella and the primary wall (Tao et al., 2009). To further investigate the effect of pollination by different parents on the deposition of lignin on stone cells in Dangshan su, TEM was used to observe DY and DW sections at different stages after flowering. Figure 6 clearly showed the lignin deposition on stone cells at different developmental stages. It could be seen that the lignin deposition process on stone cells of DY and DW was no significant difference.

The process of lignin deposition and stone cell development was generally as follows: 1. 23 DAF, lignin was deposited from the corner of cell (CC) in dispersed strip granules. With deposition of cellulose microfibrils (CMfs) along the perimeter of cell and oriented arrangement, lignification process began to expand along the CC of primary wall to other regions of CML and S layer (Figures 5A1, B1); 2. 31-39 DAF, granular lignin was in homogeneous deposited on S_1 layer oriented CMfs, starting from the inner side of S_1 (Figures 5A2-A3, B2-B3); 3. When the fruit developed to 47-55 DAF, S_2 CMfs were also arranged along the circumference of the cell and stratified. At this time, S_1 and S_2 had a relatively obvious junction (S_1L), where the

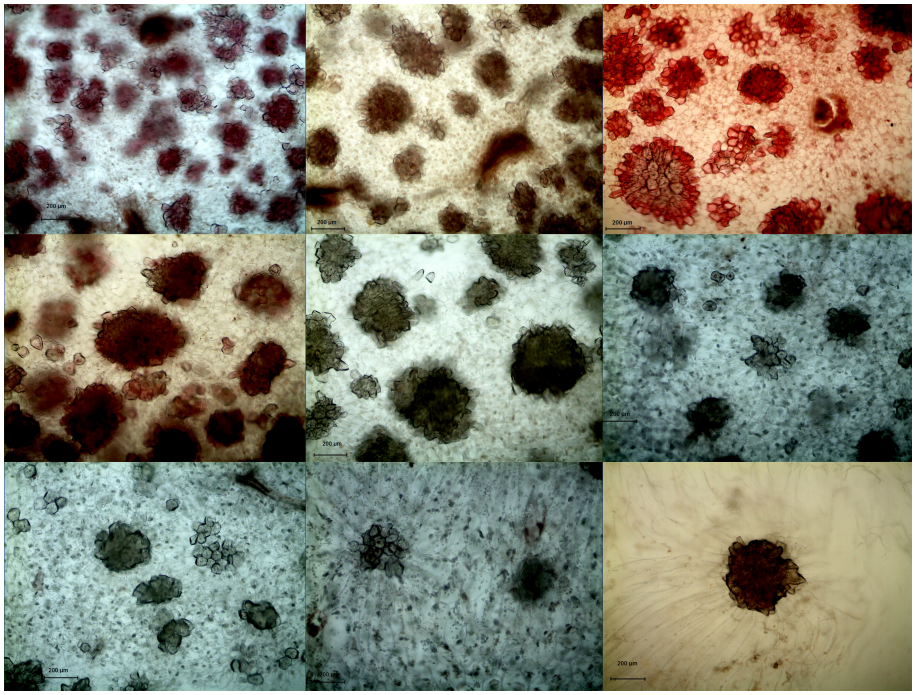


FIGURE 2
Microscopy of stone cell cluster development in DY(100×) A - I: 15d、23d、31d、39d、47d、55d、63d、87d、145d, respectively.

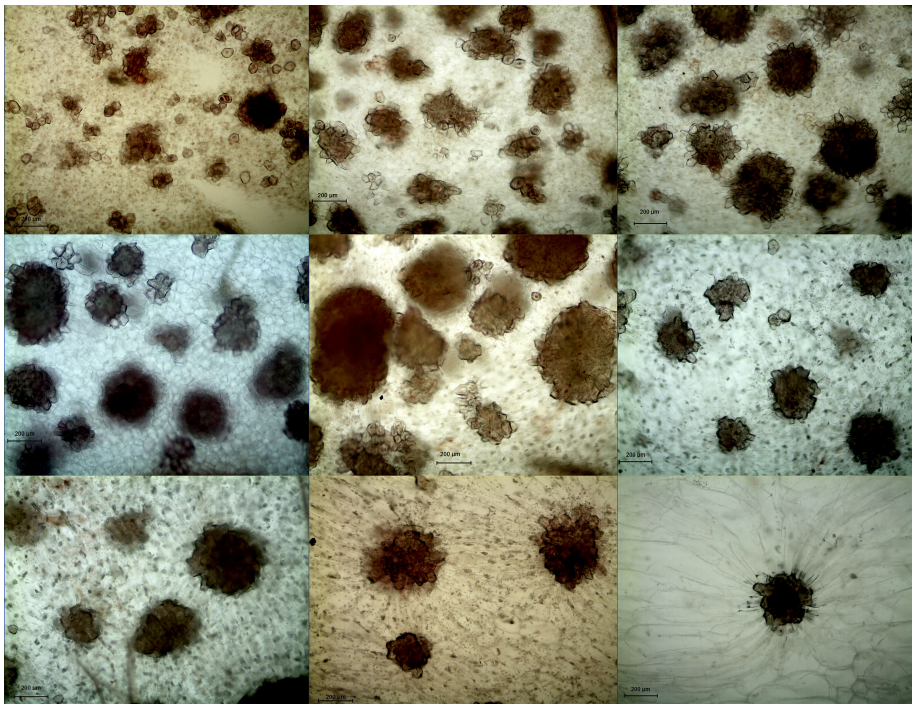


FIGURE 3
Microscopy of stone cell cluster development in DW (100 x) A - I: 15d、23d、31d、39d、47d、55d、63d、87d、145d, respectively.

staining was deep, indicating that there was more lignin deposition. Lignin particles were inhomogeneous deposited along the CMfs at S_1L, and lignin deposition started from the CMfs in the inner layer of S_2 (Figures 5A4-A5, B4-B5); 4. As lignification proceeded, the CMfs of the secondary wall S_3 remain oriented along the circumference of the cell and the lignin particles continue to be deposited unevenly along the CMfs. As lignin particles began to deposit in S_3, lignin was rapidly deposited throughout the secondary wall, forming alternating light and dark streaks, with the light streaks being CMfs and the dark streaks being lignin-deposited particles, until they filled the whole cell lumen (As shown by the short arrows in A6, B6).

Although the development processes of DY and DW stone cells were basically the same, there were differences in the compactness of the secondary wall of stone cells in different developmental stages. The compactness of the secondary wall of DY lignin deposition was higher than that of DW. Statistical analysis of the micrographs of lignin deposition at different developmental stages of DY and DW showed that the width of the secondary parietal layer of lignin deposition in DW fruits was larger than that in DY at all developmental stages, especially at 47 DAF, where the former was 2.34 times greater than the latter (Figure 6). This may be due to the different ratios of lignin G/S monomers and bond structures in pear fruit after pollination by different parents. At the same time, due to the different compactness of lignin deposition in the secondary parietal layer, the resistance of the SSCs formed by PCs to the swelling pressure caused by PCs expansion was different, which might explain the different changes in the area of SCCs in pears pollinated by different parents during late fruit development. This might also be one of the important factors causing the difference in the DP of DY and DW SCCs.

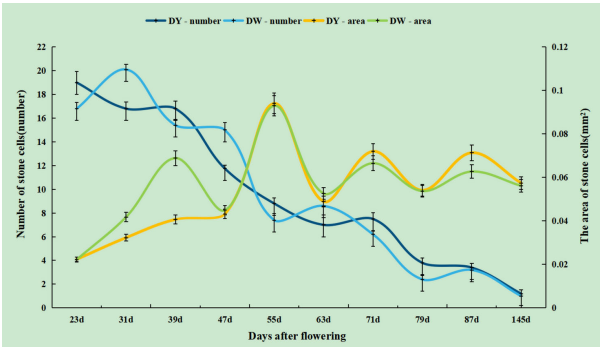


FIGURE 4
The number and area of stone cell clusters during different developmental stages of DY and DW.

10.3389/fpls.2023.1093661

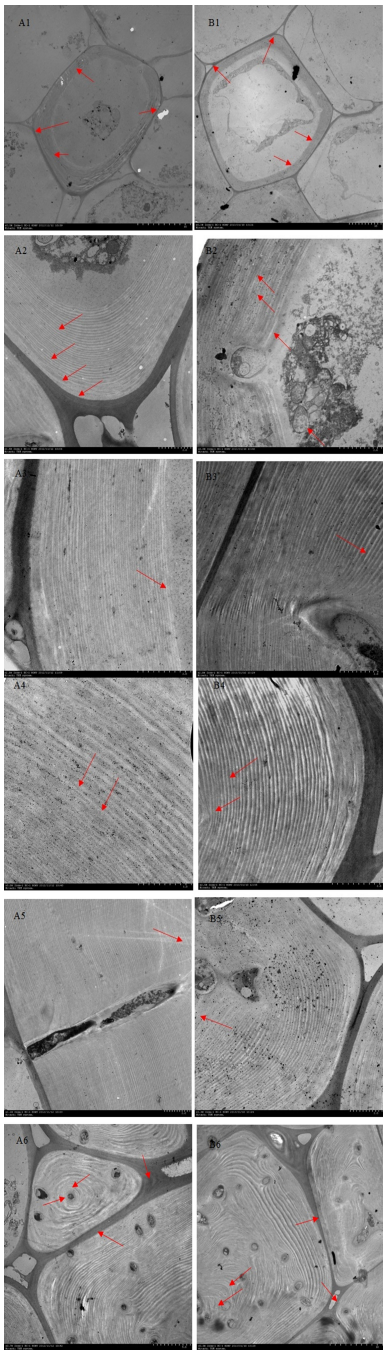


FIGURE 5
Ultramicroscopy of lignin deposition during stone cell development of DY and DW. Note: **(A1 – A6)** 23d, 31d, 39d, 47d, 55 d and 63 days after flowering of DY, respectively; **(B1 – B6)** 23d, 31d, 39d, 47d, 55d and 63 days after flowering of DW, respectively. **(A1,B1)** Lignin was unevenly deposited along the primary walls fine particles at the corner of the primary cell wall; **(A2-A3,B2-B3)** Deposition initiates from the inside of the S1 layer to S2 layer; **(A4-A5, B4-B5)** Lignin particles were deposited unevenly along the inside of each micro-fibril in every S_2 layer; **(A6,B6)**Lignin was fully occupied the cell cavity; CC: cell corner; CIL: composite intercellular layer; S_1: S_1 layer of the secondary wall; S_1L: layer between S_1 and S_2; S_2: S_2 layer of the secondary wall; S_2L: layer between S_2 and S_3; S_3: S_3 layer of the secondary wall.

3.4 Ultramicroscopic observation of stone cell pit in pear

The structural units of the primary and secondary walls of plant cells were mainly bundles of cellulose molecules (Lopez and Barclay, 2017). These bundles of cellulose molecules were bonded in layers by hydrogen bonds at certain angles and directions. The thin parts that form at certain points were called pits. The protoplasmic filaments that connected two adjacent cells to each other through the stria were called plasmodesmata, which were the bridge between cells for direct material and information contact (Sasaki et al., 2017; Chukhchin et al., 2021). In order to further investigated the changes in pits during lignification of stone cells and the effect of pits on the initiation of lignin processes in parenchymal cells, we carried out ultramicroscopic observations of the changes in pits in stone cells of DY and DW (Figure 7).

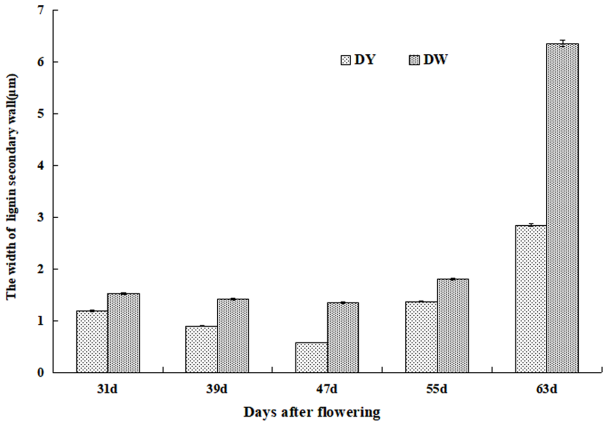


FIGURE 6
The width of Secondary wall layer of DY and DW.

As can be seen from A7 and B7, the pits of stone cells in pear fruit are mainly single pits, and two adjacent cell pits are connected by plasmodesmata in the primary cell wall. There was a large and strong presence of highly electron-dense material in the area around the pits and within the pits. By the direction of transport of these high electron-dense materials, it can be assumed that the pits on pear stone cells mainly transports degraded material from PCs that are beginning to lignify out of the cell, unlike the role of the pits in other plants, which mainly provides a water transport channel for cells with secondary walls (Sano et al., 2011).

As can be seen from A8 and B8, the PCs surrounding the stone cells also began to lignify if they were connected to the stone cells by pits, whereas the PCs not connected to the stone cells by pits did not lignify. As in A8, both 'a' and 'b' cells were PCs that were beginning to lignify, but 'c' cell was still PC. There were two pairs of pits connected between 'a' and 'b' (indicated by red arrows), but in 'a' and 'c ' and 'b' and 'c' were not connected by any pits. Also in B8, 'd', 'e' and 'f' cells were lignified PCs, while 'h ' cells and 'g' cells were PCs that had not begun to lignify. Pit pairs were connected between 'd' and 'e' and between 'e' and 'f' (red arrows indicated by), but there were no pits between 'd' and 'h', 'f' and 'h' or no pairs were found between 'd' and 'h', 'f' and 'h', or 'f' and 'g'. From this it could be seen that PCs surrounding a lignified PC, as long as there was a pits connection with the lignified PC, that PC also began to lignify to form a stone cell primordium. However, the PCs that were not connected to the PCs at the beginning of lignification were still PCs without lignification. Thus, it could be speculated that lignified PCs transported some signaling molecules that could initiate ligninization process of PCs to neighboring PCs through the pits, thus initiating lignification process of surrounding PCs. However, what such signaling molecules are remains to be further studied.

4 Discussion

Wiesner staining of cross-sections of DY and DW pear fruit, combined with SEM, revealed that at 47 - 63 DAF, DW had more SCCs than DY and fewer than DY at other times. The change trend of the DP of SCCs increased first and then decreased, reaching the maximum value at 55 DAF. The trends in the DP and number of SCCs in DY and DW were basically the same, but there were differences in content, indicating that pollination by different parents mainly affected the DP and number of SCCs in pear fruit, but did not change the stone cell development process. Therefore, it could be speculated that the number and DP of SCCs in pear fruit were quantitative traits controlled by multiple genes, which was consistent with the results of previous studies(Zhang et al., 2021).

The SCC was a group of lignified PCs, and how the "lignified" PCs "shrink" had been postulated to be due to the breakdown of the large SCCs into individual SCCs by laccase, pectinase and cellulase, which in turn were degraded by enzymes, resulting in a reduction in the DP of the SCCs at a later stage of development (Xue et al., 2019; Wang et al., 2020). However, Lee et al. suggested that the reduction in the number and DP of SCCs in late development was not due to enzymatic breakdown, but rather to the fact that SCCs form slower than the rate of fruit expansion and are 'squeezed and diluted' by PCs (Lee et al., 2006).

The stone cell development and lignin deposition at different developmental stages of DY and DW were observed by TEM and found that the processes were largely consistent. However, there were significant differences in the compactness of the lignin-deposited secondary wall wall layers, with those of DW significantly wider than those of DY. In the studies of bamboo and wood, they were found that the compactness the secondary parietal layer, the higher the stability and compressive resistance (Murphy and Alvin, 1997; Suzuki and Itoh, 2001). Therefore, it was speculated that at the late stage of fruit development, due to the different pollination fathers, the compactness of the secondary wall of stone cells was different, which led to the different ability of the stone cell to resist the expansion pressure of PCs, which led to the different changes in the SCCs DP of DY and DW, and thus affecting the DP of SCCs.

In this study, we found that the pit was mainly in simple pit pair, with the pit pairs of adjacent cells connected by plasmodesmata at the primary cell wall, which is consistent with the type of pits and the way they are connected in the cell walls of plants such as Wheat, Eucommia and Mao bamboo (Kristensen et al., 2008; Fromm, 2013; Lian et al., 2019). By analyzing the direction of transport of highly

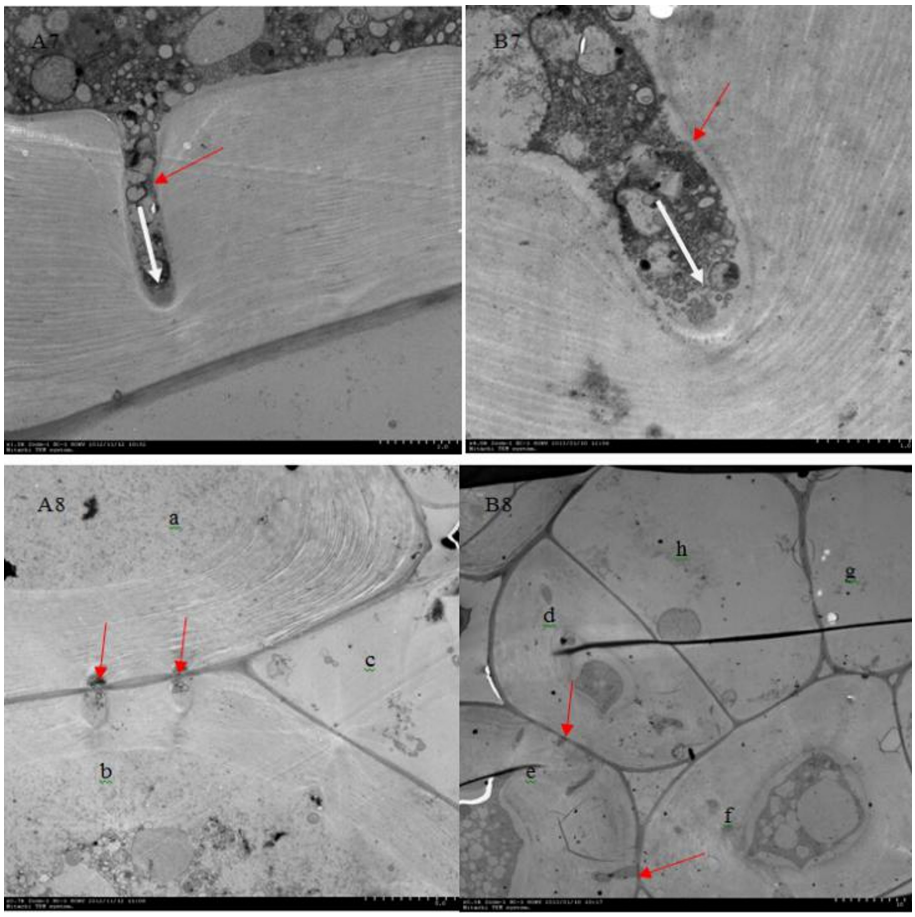


FIGURE 7
Ultramicroscopy of pits on pear stone cells of DY and DW. **(A7, A8)** was DY, **(B7, B8)** was DW. **(A7, B7)** Pits on the stone cells (short red arrow), the direction of material transport was from the intracellular to extracellular (short white arrow). **(A8)**'a' cell and 'b' cell were lignified parenchyma cells, 'c' cell was still parenchyma cell. There were two pit pairs between the 'a' cell and 'b' cell (short red arrow), but no pit pair between 'a' cell and 'c' cell or between 'b' cell and 'c' cell. **(B8)** 'd' cell, 'e' cell and 'f' cell were lignified parenchyma cells, 'h' cell an 'g' cell were just parenchyma cells. There was pit pairs between 'd' cell and 'e' cell as well as between 'e' cell and 'f' cell (short red arrow), but no pit pairs between 'd' cell and 'g' cell or between 'd' cell and 'h' cell or between 'f' cell.

electron-dense material within the pits, it was hypothesized that the role of the pits on pear stone cells was to transport degraded material out of the cell from PCs which were beginning to lignify, and that the transported materials might contain some kind of signalling molecules that could initiate the process of lignification in PCs, thus accelerating the process of lignification in PCs surrounding the stone cells. It had been speculated that the substance in pear stone cells may be calcium ions (Lee et al., 2006; Tao et al., 2009), and it had also been speculated in moso bamboo and wood that the signaling molecules may be H_2O_2 (Lian et al., 2020), but neither could give sufficient evidence, so the specific substance of the signaling molecules need to be further studied.

Therefore, a better understanding of the effects of pollination by different parents on pear stone cell development and lignin deposition. It can provide a scientific basis for clarifying the effect of different parental pollination on pear stone cell development, and thus having great significance for pear breeding.

5 Conclusion

This study proved that the formation process of DY and DW SCC was consistent, but the SCC number and DP in DY was higher than that in DW. All of the ultramicroscopy results regarding stone cell development and lignin deposition in Dangshan Su pears are summarized in Figure 8: The lignification process of DY and DW were all from corner to rest regions of the compound middle lamella and the secondary wall, with lignin particles deposited along the cellulose microfibrils. They were alternatively arranged until they filled up the whole cell cavity to culminate in the formation of stone cells. However, the compactness of the wall layer of cell wall DY was significantly higher than that of cell wall DW. Therefore, DY SCC had a higher ability to resist the expansion pressure of PC. We also found that the pit of stone cells was predominantly single pit pair, they transported degraded material from the PC that were beginning to lignify out of the cells.

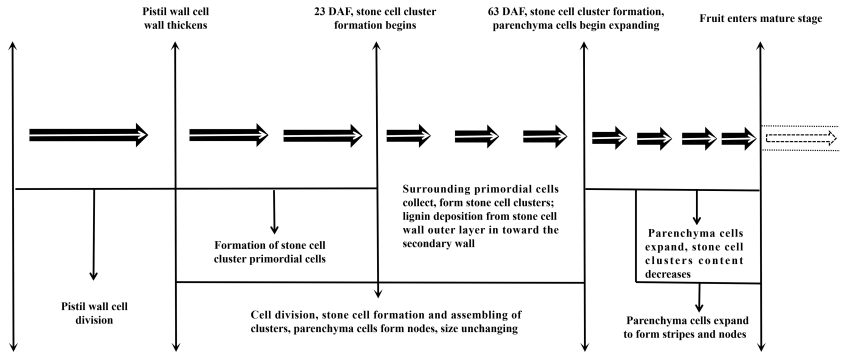


FIGURE 8
Stone cell Development in pear.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

CY and NZ performed the experiments and wrote the paper. CX contributed reagents/materials/analysis tools. CY, NZ and QJ analyzed the data. CY, YC and YQ discussed and analyzed the results. YC conceived and designed the experiments. All authors read and approved the final manuscript. CY and NZ have contributed equally to this work. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31171944, 31640068), Anhui Provincial Natural Science Foundation (Grant No. 2019B319), Earmarked Fund for Anhui Science and Technology Major Project (202003b06020016).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Cai, Y., Li, G., Nie, J., Lin, Y., Nie, F., Zhang, J., et al. (2010). Study of the structure and biosynthetic pathway of lignin in stone cells of pear. *Scientia Hortic.* 125 (3), 374–379. doi: 10.1016/j.scienta.2010.04.029

Cheng, X., Cai, Y., and Zhang, J. (2019). Stone cell development in pear. *Pear Genome*, 201–225. doi: 10.1007/978-3-030-11048-2_11

Cheng, X., Yan, C., Zhang, J., Ma, C., Li, S., Jin, Q., et al. (2017). The effect of different pollination on the expression of dangshan su pear MicroRNA. *BioMed. Res. Int.* 2017, 2794040. doi: 10.1155/2017/2794040

Cheng, X., Zhang, J., Wang, H., Chen, T., Li, G., Yan, C., et al. (2020). Effects of metaxenia on stone cell formation in pear (Pyrus bretschneideri) based on transcriptomic analysis and functional characterization of the lignin-related gene PbC4H2. *Forests* 11 (1), 53. doi: 10.3390/f11010053

Chukhchin, D. G., Vashukova, K., and Novozhilov, E. (2021). Bordered pit formation in cell walls of spruce tracheids. *Plants* 10 (9), 1968. doi: 10.3390/plants10091968

Clifford, M. N. (1974). Specificity of acidic phloroglucinol reagents. *J. Chromatogr. A* 94, 321–324. doi: 10.1016/S0021-9673(01)92389-1

Day, A., Ruel, K., Neutelings, G., Crônier, D., David, H., Hawkins, S., et al. (2005). Lignification in the flax stem: evidence for an unusual lignin in bast fibers. *Planta* 222 (2), 234–245. doi: 10.1007/s00425-005-1537-1

Escamez, S., Bollhöner, B., and Tuominen, H. (2017). "Quick histochemical staining methods to detect cell death in xylem elements of plant tissues," in *Xylem: Methods and protocols.* Eds. M. de Lucas and J. P. Etchhells (New York, NY: Springer New York), 27–36.

Fromm, J. (2013). "Xylem development in trees: from cambial divisions to mature wood cells," in *Cellular aspects of wood formation* (Springer), 3–39. doi: 10.1007/978-3-642-36491-4_1

Guo, Y., Zhou, J., Wen, J., Sun, G., and Sun, Y. (2015). Structural transformations of triploid of populus tomentosa carr. lignin during auto-catalyzed ethanol organosolv pretreatment. *Ind. Crops Products* 76, 522–529. doi: 10.1016/j.indcrop.2015.06.020

Jin, Q., Yan, C., Qiu, J., Zhang, N., Lin, Y., and Cai, Y. (2013). Structural characterization and deposition of stone cell lignin in dangshan su pear. *Scientia Hortic.* 155, 123–130. doi: 10.1016/j.scienta.2013.03.020

Kristensen, J. B., Thygesen, L. G., Felby, C., Jørgensen, H., and Elder, T. (2008). Cell-wall structural changes in wheat straw pretreated for bioethanol production. *Biotechnol. Biofuels* 1 (1), 1–9. doi: 10.1186/1754-6834-1-5

Lee, S.-H., Choi, J.-H., Kim, W.-S., Han, T.-H., Park, Y.-S., and Gemma, H. (2006). Effect of soil water stress on the development of stone cells in pear (Pyrus pyrifolia cv.'Niitaka') flesh. *Scientia Hortic.* 110 (3), 247–253. doi: 10.1016/j.scienta.2006.07.012

Lian, C., Liu, R., Xiufang, C., Zhang, S., Luo, J., Yang, S., et al. (2019). Characterization of the pits in parenchyma cells of the moso bamboo [Phyllostachys edulis (Carr.) j. houz.] culm. *Holzforschung* 73 (7), 629–636. doi: 10.1515/hf-2018-0236

Lian, C., Liu, R., Zhang, S., Yuan, J., Luo, J., Yang, F., et al. (2020). Ultrastructure of parenchyma cell wall in bamboo (Phyllostachys edulis) culms. *Cellulose* 27 (13), 7321–7329. doi: 10.1007/s10570-020-03265-9

Li, S., Su, X., Jin, Q., Li, G., Sun, Y., Abdullah, M., et al. (2018). iTRAQ-based identification of proteins related to lignin synthesis in the pear pollinated with pollen from different varieties. *Molecules* 23 (3), 548. doi: 10.3390/molecules23030548

Lopez, F., and Barclay, G. (2017). "Plant anatomy and physiology," in *Pharmacognosy* (Elsevier), 45–60. doi: 10.1016/B978-0-12-802104-0.00004-4

Mumtaz, S., Ganai, N., Hamid, M., Ahad, S., Ashraf, S., Javid, R., et al. (2020). Effect of pollen source on some fruit characteristics of exotic pear (Pyrus communis l.) cv. carmen and abate fetel. *J. Pharmacognosy Phytochem.* 9 (6), 120–123. doi: 10.22271/phyto.2020.v9.i6b.12869

Murphy, R., and Alvin, K. (1997). Fibre maturation in the bamboo gigantochloa scortechinii. *IAWA J.* 18 (2), 147–156. doi: 10.1163/22941932-90001476

Nie, J.-q., Cai, Y.-p., Zhang, S.-h., Lin, Y., Xu, Y.-l., and Zhang, J.-y. (2009). The anatomic study on relationship of stone cells and parenchyma cells during fruit development of pyrus bretschneideri. *Acta Hortic. Sin.* 8, 019. doi: 10.16420/j.issn.0513-353x.2009.08.019

Sano, Y., Morris, H., Shimada, H., Ronse De Craene, L. P., and Jansen, S. (2011). Anatomical features associated with water transport in imperforate tracheary elements of vessel-bearing angiosperms. *Ann. Bot.* 107 (6), 953–964. doi: 10.1093/aob/mcr042

Sasaki, T., Fukuda, H., and Oda, Y. (2017). CORTICAL MICROTUBULE DISORDERING1 is required for secondary cell wall patterning in xylem vessels. *Plant Cell* 29 (12), 3123–3139. doi: 10.1105/tpc.17.00663

Shutian Tao, S. K., Zhang, H., and Zhang, S. (2009). Anatomy, ultrastructure and lignin distribution of stone cells in two pyrus species. *Plant Sci.* 176 (2009), 7. doi: 10.1016/j.plantsci.2008.12.011

Suzuki, K., and Itoh, T. (2001). The changes in cell wall architecture during lignification of bamboo, phyllostachys aurea Carr. *Trees* 15 (3), 137–147. doi: 10.1007/s004680000084

Tao, S., Khanizadeh, S., Zhang, H., and Zhang, S. (2009). Anatomy, ultrastructure and lignin distribution of stone cells in two pyrus species. *Plant Sci.* 176 (3), 413–419. doi: 10.1016/j.plantsci.2008.12.011

Wang, X., Liu, S., Liu, C., Liu, Y., Lu, X., Du, G., et al. (2020). Biochemical characterization and expression analysis of lignification in two pear (Pyrus ussuriensis maxim.) varieties with contrasting stone cell content. *Protoplasma* 257 (1), 261–274. doi: 10.1007/s00709-019-01434-7

Xue, C., Yao, J. L., Qin, M. F., Zhang, M. Y., Allan, A. C., Wang, D. F., et al. (2019). PbrmiR397a regulates lignification during stone cell development in pear fruit. *Plant Biotechnol. J.* 17 (1), 103–117. doi: 10.1111/pbi.12950

Xu, F., Zhong, X. C., Sun, R. C., and Lu, Q. (2006). Anatomy, ultrastructure and lignin distribution in cell wall of caragana korshinskii. *Ind. Crops Products* 24 (2), 186–193. doi: 10.1016/j.indcrop.2006.04.002

Xu, Y, Zhang Jinyun, G. Z., Xingkai, Y., and Bing, S.. (2009). Effects of pollination with various pollination varieties and pollen amount on fruit calyx ofDangshan Suli pear. *Journal of Anhui Agricultural University* 36 (1), 1–6. doi: 10.1038/s41467-021-21378-y

Yan, C., Yin, M., Zhang, N., Jin, Q., Fang, Z., Lin, Y., et al. (2014). Stone cell distribution and lignin structure in various pear varieties. *Scientia Hortic.* 174, 142–150. doi: 10.1016/j.scienta.2014.05.018

Yin, Y., Jiang, X., and Qu, C. (2003). Dynamic changes of lignin deposition in secondary xylem cell wall during secondary xylem differentiation in populus tomentosa carr. *J. Chin. Electron Microscopy Soc.* 23 (6), 663–669. doi: 10.3969/j.issn.1000-6281.2004.06.016

Zhang, M. Y., Xue, C., Hu, H., Li, J., Xue, Y., Wang, R., et al. (2021). Genome-wide association studies provide insights into the genetic determination of fruit traits of pear. *Nat. Commun.* 12 (1), 1144. doi: 10.1038/s41467-021-21378-y

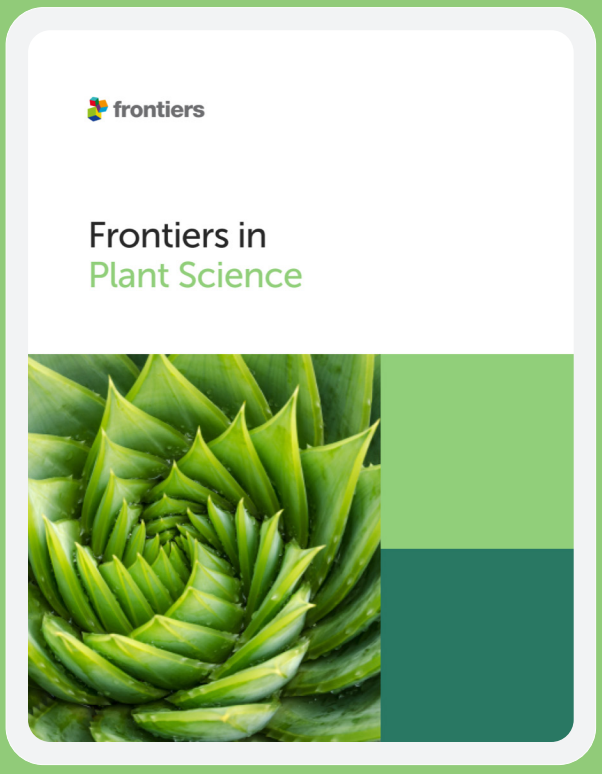
Frontiers in Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact