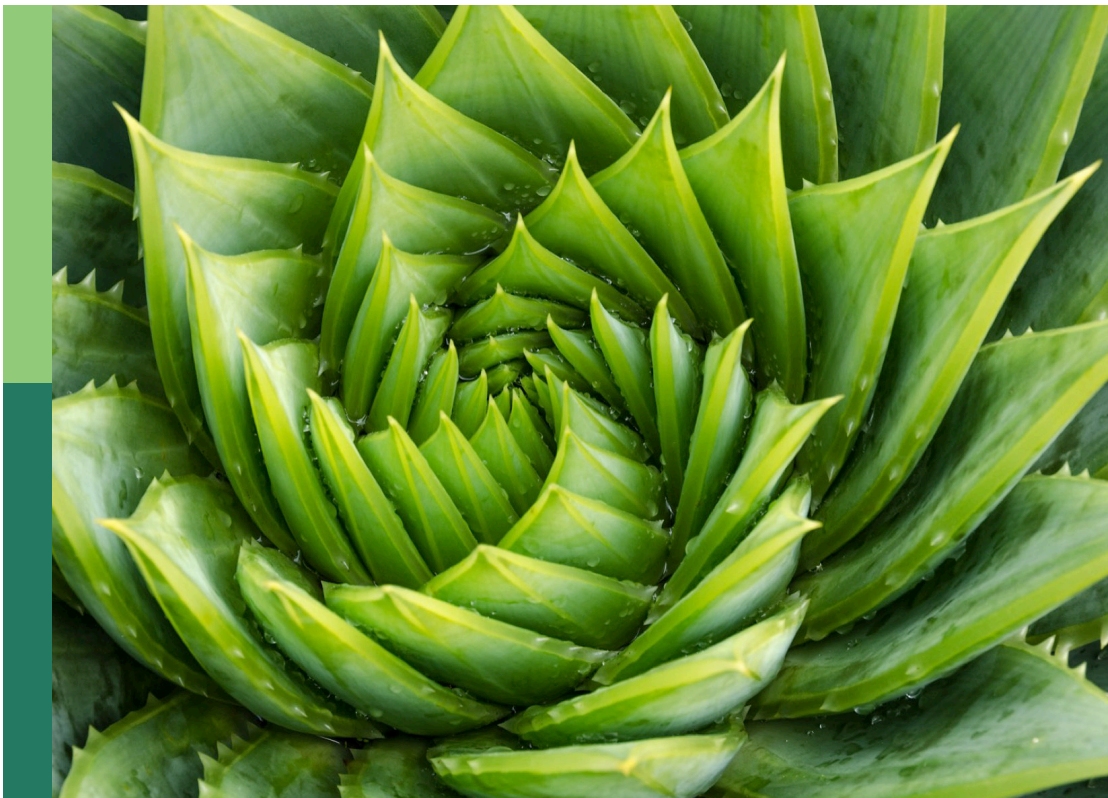
frontiers | Research Topics

Flowering time control in agricultural and horticultural crops

Edited by
Liang Wu, Leo Marcelis, Fanjiang Kong and Yang Zhu

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-83252-016-1
DOI 10.3389/978-2-83252-016-1

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Flowering time control in agricultural and horticultural crops

Topic editors
Liang Wu — Zhejiang University, China
Leo Marcelis — Wageningen University and Research, Netherlands
Fanjiang Kong — Guangzhou University, China
Yang Zhu — Zhejiang University, China

Citation
Wu, L., Marcelis, L., Kong, F., Zhu, Y., eds. (2023). *Flowering time control in agricultural and horticultural crops*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83252-016-1

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 09 February 2023
DOI 10.3389/fpls.2023.1116197

OPEN ACCESS

EDITED AND REVIEWED BY
Ildikó Karsai,
Centre for Agricultural Research, Hungary

*CORRESPONDENCE
Liang Wu
liangwu@zju.edu.cn

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 05 December 2022
ACCEPTED 02 February 2023
PUBLISHED 09 February 2023

CITATION
Wu L, Marcelis LFM, Kong F and Zhu Y (2023) Editorial: Flowering time control in agricultural and horticultural crops. *Front. Plant Sci.* 14:1116197. doi: 10.3389/fpls.2023.1116197

Editorial: Flowering time control in agricultural and horticultural crops

Liang Wu[1,2]*, Leo F.M. Marcelis[3], Fanjiang Kong[4] and Yang Zhu[2]

[1]Hainan Yazhou Bay Seed Laboratory, Hainan Institute, Zhejiang University, Sanya, Hainan, China, [2]College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, Zhejiang, China, [3]Horticulture and Product Physiology, Department of Plant Sciences, Wageningen University, Wageningen, Netherlands, [4]School of Life Sciences, Guangzhou University, Guangzhou, China

KEYWORDS

flowering regulation, photoperiod, vernalization, circadian clock, seasonal flowering, yield potential

Editorial on the Research Topic
Flowering time control in agricultural and horticultural crops

Flowering time regulation in higher plants is a crucial developmental phase change controlled by environmental conditions and developmental signaling. The complexity of this regulation divides plant growth into vegetative and reproductive stages, which are processed by an extensive network of floral signaling pathways. Decades of research on flowering time control have discovered several comprehensive systems that plants have evolved to trigger the floral transition. Factors such as day length, temperature, florigen movement, vernalization, and the juvenile phase transition play regulatory roles in flowering control. This Research Topic aims to illustrate advanced findings in flowering time control in agricultural and horticultural plants. Horticultural crops provide special food resources for people, such as fruits, vegetables and tea, and ornamentals add beauty and interest to a garden, thus understanding the nature of their reproductive transitions will reinforce the quality and thereby increasing the income and the prosperity to benefit mankind.

Among the regulatory factors, photoperiod has been considered a central environmental factor for flowering time control in plants. Based on the molecular mechanisms of flowering time, short-day plants (SDPs) and long-day plants (LDPs) have been gradually investigated, which offers an opportunity to explain the changes in day length observed by the two types of plants (Song et al., 2018). In LDP, such as Arabidopsis, CONSTANS (CO) is a pivotal hub integrating numerous internal and external signals for inducing photoperiodic flowering. In SDP, such as rice, the homolog of CO, Hd1, on the other hand, promotes and prevents flowering under SD and LD, respectively (Shim et al., 2017). Many SDPs, such as rice and soybean, have a critical threshold for day length and can even detect changes of 15 minutes for flowering decisions. To better understand the critical day length regulation and the conflict between SDPs and LDPs through photoperiod, Lv et al. conducted a study using next-generation sequencing (NGS)-based bulked segregant analysis (BSA) to map

quantitative genes controlling the long-juvenile (LJ) trait in soybean. The LJ trait has been introduced into soybean cultivars to increase yield in tropical environments. The authors identified two genomic regions on scaffold 32 and chromosome 18 harboring loci LJ32 and LJ18, respectively, regulating LJ trait controlling soybean flowering. Their findings will enhance our understanding of the molecular mechanisms underlying the LJ trait and provide useful genetic resources for soybean molecular breeding in tropical regions. In another study, Khan et al. characterized the natural variations in CO family genes and their association with flowering time and maturity in soybean regarding the whole genomic region involving conserved and mutated genes. The authors reported that haplotypes exhibited natural divergence associated with flowering dates and soybean maturity in adapting to diverse environments.

The onset of flowering in plants is precisely controlled by extensive environmental factors and internal molecular networks, in which FLOWERING LOCUS T (FT) is a key flowering integrator (Cho et al., 2017). Yuan et al. studied FT homologs in the soybean genome, diversifying the soybean flowering pathway through characterizing *GmFT3a*, *GmFT2a*, and *GmFT5a*, which are located on the same chromosome as the flowering promoters and act as flowering promoters in the non-inductive photoperiod in soybean. The authors suggested that *GmFT3a* provides an opportunity to slightly promote the flowering time of soybean varieties, which helps retain the yield and agronomic traits of an elite variety with the extension of its adaptive regions (Qi et al., 2021).

The circadian clock and related genes are another internal timing mechanism that allows plants to make decisions in accordance with the environmental conditions. Gong et al. examined the regulatory roles of the circadian clock genes on growth and development of crop species on yield-related traits characterizing CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1) homologue in wheat. Their results provide novel insights into a circadian-mediated mechanism for expressing genes in wheat to coordinate photosynthetic and metabolic activities, leading to optimal growth and development. The authors investigated the tissue-specific expression pattern of TaCCA1 genes in wheat at ZT3 grown under LD conditions. All three TaCCA1 genes exhibited analogous expression patterns with a strong accumulation in green tissues than the other tissues, which were mostly the same as those observed in LHY/CCA1 of Arabidopsis and OsCCA1 of rice (Lu et al., 2009; Sun et al., 2021).

SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1) encodes a MADS-box protein that plays regulatory roles in integrating multiple flowering signals for floral transition and reproductive development in Arabidopsis. Ma and Yan investigated how cotton responds to environmental cues to adjust flowering time in achieving the reproductive success. Cotton is cultivated worldwide due to its broader adaptation to environment and successful breeding for early-matured varieties. This study highlighted the roles of *SOC1* for the integration of endogenous and exogenous signals in plants to maximize reproduction, demonstrating that GhSOC1s may evolve divergently, respond differently to light and temperature, and act cooperatively to promote floral transition in tetraploid cotton.

Transcription factors, besides other endogenous developmental signals, are also involved in the flowering time regulation in plants. WRKY is a large family of transcription factors known for various functions ranging from stress resistance to plant growth and development (Li et al., 2016). Khan et al. studied Chrysanthemum, a well-known ornamental plant with multiple uses, examined the WRKY family and observed a total 138 genes which were classified into III groups. Group III of *C. lavandulifolium* contains 53 members, which is larger than group III of Arabidopsis. AuR and GRE-responsive *cis*-acting elements were located in the promoter region of WRKY members, which are important for plant development and flowering induction. This research provides a basis to study the role of WRKY genes in developing ornamental plants, especially in flowering traits.

Trevaskis et al. wrote a comprehensive review discussing Oat (*Avena sativa*) seasonal flowering behavior as a key contributor in the successful cultivation of oat. Oat is a vernalization-responsive LD plant that flowers after winter as daytime lengthens in spring. Variation in both vernalization and day length requirements broadens the adaptation of oat and has been used for breeding modern cultivars with seasonal flowering behaviors suited to different regions, sowing dates, and farming practices. The authors examined the importance of variation in oat phenology for crop adaptation and outlined strategies to advance our understandings of the genetic basis of oat phenology. Their research emphasized the molecular basis of oat phenology to resolve the contribution of individual genes to crop performance by developing oat genetic resources, such as near-isogenic lines and genetically modified new oat varieties.

Higher and more stable crop yields are the main targets for cereal breeders. It is expected that vernalization requirements of current cultivars can be desynchronized with the environment's vernalizing potential as the winters in temperate areas will become warmer (Wu et al., 2017). To honor the promise of increased yield potentials, Fernández-Calleja et al. studied hybrid barley phenology deployed in new cultivars. In their study, hybrid combinations extend the available catalog of genetic responses to vernalization, opening new possibilities for optimizing the phenology to specific areas using hybrids. Hybrids can show a more nuanced response to insufficient vernalization than inbred lines proposing new options to manage flowering time based on specific alleles and, particularly, the duration of developmental phases that build yield potential in hybrid barley, suggesting that this strategy could be used in other crops to provide future food security.

RNA sequencing (RNA-Seq) is a powerful tool to examine the continuously changing cellular transcriptome under distinct conditions, therefore, Dong et al. conducted RNA-seq to elucidate the expression of light-related regulatory genes under SD photoperiod inducement of adzuki bean flowering, providing an important theoretical basis for accelerated breeding programs. Their study provides a deep understanding of the molecular mechanisms of adzuki bean flowering in response to SD photoperiod, which laid a foundation for the functional verification of genes delivering an important reference for the molecular breeding of adzuki beans for further studies.

Together, these recently published articles provide a broad overview of the key roles of flowering time control and related genes in agricultural and horticultural crops. This Research Topic highlights innovative and

emerging areas in the discipline and opens new routes of discovery that will inspire researchers with extensive research interests.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (LZ21C130002).

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

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

Li, W., Wang, H., and Yu, D. (2016). Arabidopsis WRKY transcription factors WRKY12 and WRKY13 oppositely regulate flowering under short-day conditions. *Mol. Plant* 9, 1492–1503. doi: 10.1016/j.molp.2016.08.003

Lu, S. X., Knowles, S. M., Andronis, C., Ong, M. S., and Tobin, E. M. (2009). CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL function synergistically in the circadian clock of arabidopsis. *Plant Physiol.* 150, 834–843. doi: 10.1104/pp.108.133272

Qi, X., Jiang, B., Wu, T., Sun, S., Wang, C., Song, W., et al. (2021). Genomic dissection of widely planted soybean cultivars leads to a new breeding strategy of crops in the post-genomic era. *Crop J.* 9, 1079–1087. doi: 10.1016/j.cj.2021.01.001

Shim, J. S., Kubota, A., and Imaizumi, T. (2017). Circadian clock and photoperiodic flowering in arabidopsis: CONSTANS is a hub for signal integration. *Plant Physiol.* 173, 5–15. doi: 10.1104/pp.16.01327

Song, Y. H., Kubota, A., Kwon, M. S., Covington, M. F., Lee, N., Taagen, E. R., et al. (2018). Molecular basis of flowering under natural long-day conditions in arabidopsis. *Nat. Plants* 4, 824–835. doi: 10.1038/s41477-018-0253-3

Sun, C., Zhang, K., Zhou, Y., Xiang, L., He, C., Zhong, C., et al. (2021). Dual function of clock component OsLHY sets critical day length for photoperiodic flowering in rice. *Plant Biotechnol. J.* 19, 1644–1657. doi: 10.1111/pbi.13580

Wu, X., Liu, H., Li, X., Tian, Y., and Mahecha, M. D. (2017). Responses of winter wheat yields to warming-mediated vernalization variations across temperate Europe. *Front. Ecol. Evol.* 5, 126. doi: 10.3389/fevo.2017.00126

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 March 2022
doi: 10.3389/fpls.2022.817544

CONSTANS Polymorphism Modulates Flowering Time and Maturity in Soybean

Mohammad Abdul Awal Khan[1†], Shouwei Zhang[1†], Reza Mohammad Emon[1,2†], Fulu Chen[1†], Wenwen Song[1], Tingting Wu[1], Shan Yuan[1], Cunxiang Wu[1], Wensheng Hou[1], Shi Sun[1], Yongfu Fu[1], Bingjun Jiang[1] and Tianfu Han[1]**

[1] MARA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [2] Plant Breeding Division, Bangladesh Institute of Nuclear Agriculture, Mymensingh, Bangladesh

CONSTANS (*CO*) plays a critical role in the photoperiodic flowering pathway. However, the function of soybean *CO* orthologs and the molecular mechanisms in regulating flowering remain largely unknown. This study characterized the natural variations in *CO* family genes and their association with flowering time and maturity in soybeans. A total of 21 soybean *CO* family genes (*GmCOL*s) were cloned and sequenced in 128 varieties covering 14 known maturity groups (MG 0000-MG X from earliest to latest maturity). Regarding the whole genomic region involving these genes, *GmCOL1*, *GmCOL3*, *GmCOL8*, *GmCOL9*, *GmCOL10,* and *GmCOL13* were conserved, and the remaining 15 genes showed genetic variation that was brought about by mutation, namely, all single-nucleotide polymorphisms (SNPs) and insertions-deletions (InDels). In addition, a few genes showed some strong linkage disequilibrium. Point mutations were found in 15 *GmCOL* genes, which can lead to changes in the potential protein structure. Early flowering and maturation were related to eight genes (*GmCOL1/3/4/8/13/15/16/19*). For flowering and maturation, 11 genes (*GmCOL2/5/6/14/20/22/23/24/25/26/28*) expressed divergent physiognomy. Haplotype analysis indicated that the haplotypes of *GmCOL5-Hap2*, *GmCOL13-Hap2/3,* and *GmCOL28-Hap2* were associated with flowering dates and soybean maturity. This study helps address the role of *GmCOL* family genes in adapting to diverse environments, particularly when it is necessary to regulate soybean flowering dates and maturity.

Keywords: soybean, *GmCOL* orthologue, natural variation, flowering time, maturity group

OPEN ACCESS

Edited by:
Fanjiang Kong,
Guangzhou University, China

Reviewed by:
Chao Fang,
Michigan State University, United States
Steven B. Cannon,
Agricultural Research Service (USDA), United States

***Correspondence:**
Bingjun Jiang
jiangbingjun@caas.cn
Tianfu Han
hantianfu@caas.cn

[†]These authors have contributed equally to this work

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

Received: *18 November 2021*
Accepted: *15 February 2022*
Published: *17 March 2022*

Citation:
Awal Khan MA, Zhang S, Emon RM, Chen F, Song W, Wu T, Yuan S, Wu C, Hou W, Sun S, Fu Y, Jiang B and Han T (2022) CONSTANS Polymorphism Modulates Flowering Time and Maturity in Soybean. Front. Plant Sci. 13:817544. doi: 10.3389/fpls.2022.817544

INTRODUCTION

Plants can adapt to different environmental conditions in response to various day lengths (photoperiods). In the photoperiodic flowering pathway, *CONSTANS* (*CO*), which is a B-box-containing transcription factor (Robson et al., 2001; Gangappa and Botto, 2014), plays a key role (Putterill et al., 1995; Suarez-Lopez et al., 2001). CO also possesses a CONSTANS, CONSTANS-LIKE, TIMING OF CAB1 (CCT) domain at its C-terminus involved in DNA binding (Strayer et al., 2000; Robson et al., 2001). CYCLING DOF FACTOR (CDF) family proteins bind to the *CO* promoter to repress its transcription in the morning (Imaizumi et al., 2005; Fornara et al., 2009). FLAVIN-BINDING, KELCH REPEAT, and F-BOX1 (FKF1) interact with GIGANTEA (GI) to degrade the CDF1 protein, which results in the elevation of *CO* mRNA (Sawa et al., 2007). CO acts in the phloem of leaf vascular tissues to activate *FLOWERING LOCUS T* (*FT*)

expression, which causes flowering under linkage disequilibrium (LD) conditions (Takada and Goto, 2003; An et al., 2004). SUPPRESSOR OF PHYTOCHROME A-105 (SPA) proteins interact with the CCT domain of CO (Laubinger et al., 2006). CO is ubiquitinated by CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) and degraded by the 26S proteasome (Jang et al., 2008; Liu L. J. et al., 2008). HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE GENES1 (HOS1) interacts with CO and participates in the degradation of CO mediated by red light (Lazaro et al., 2012, 2015). Nucleoporin96 (Nup96) interacts with HOS1 to gate CO protein levels (Cheng et al., 2020). CO protein is stable under the light in the evening but degraded in the morning or in the dark under LD conditions (Valverde et al., 2004). All of the abovementioned comments are based on studies in *Arabidopsis*.

As a short-day plant (SDP) distributed over a vast range of latitudes, soybean is characterized as having 14 maturity groups from MG 0000 to MG X (Liu et al., 2017; Jiang et al., 2019). There are hundreds of genetic loci associated with flowering time and maturity in soybean, among which *E1* (Xia et al., 2012), *E2* (Watanabe et al., 2011), *E3* (Watanabe et al., 2009), *E4* (Liu B. et al., 2008), *E6* (Fang C. et al., 2021), *E9* (Kong et al., 2014), *E10* (Samanfar et al., 2017), *E11* (Wang et al., 2019), *J* (Lu et al., 2017; Yue et al., 2017) and *Tof11/GmPRR37/GmPRR3b* (Li et al., 2020; Lu et al., 2020; Wang et al., 2020), *GmTof16* (Dong et al., 2021b), *GmFUL* (Dong et al., 2021a; Sun et al., 2021; Yue et al., 2021), and *GmLUX* (Bu et al., 2021; Fang X. et al., 2021) are molecularly identified. There are 28 *CO* orthologs in soybean (Fan et al., 2014), but the functions of most of the CO orthologs remain uncharacterized. It has been shown that *GmCOL1a* and *GmCOL1b* can fully complement the late-flowering phenotype of *the co1* mutant in *Arabidopsis* (Wu et al., 2014). In contrast, *GmCOL1a* and *GmCOL1b* repressed flowering in soybean under long-day conditions (Cao et al., 2015; Wu et al., 2019).

Natural variation is associated with photoperiodic flowering and adaptation in different species (Alonso-Blanco et al., 2005; Balasubramanian et al., 2006; Rosas et al., 2014; Li et al., 2017; Lu et al., 2017; Bao et al., 2019; Jiang et al., 2019; Wu et al., 2020). In this study, 128 soybean varieties were selected and planted, covering all 14 maturity groups from MG 0000 to MG X with a continuous distribution in maturity groups (Jiang et al., 2019). Due to the short duration of the project, 21 of the 28 soybean *COL* genes (Fan et al., 2014) were sequenced, and their sequence polymorphisms were analyzed. Furthermore, we studied the haplotypes of these soybean *GmCOL* genes to discover their natural variations in association with flowering time and maturity. These results suggested that some natural variations in 21 soybean *GmCOL* genes were associated with flowering date and maturity.

MATERIALS AND METHODS

Plant Materials

Soybean varieties covering all 14 maturity groups (MG 0000-MG X) were assessed in this study (Jiang et al., 2019). In this study, 64 Chinese and Russian soybean varieties and 64 North American maturity group standard varieties were included, covering MG 0000-MG X, for a total of 128 accessions analyzed (**Table 1**).

Investigation of Flowering and Maturity Dates

The soybeans were planted in soil in 10-L pots and grown under natural conditions in Haidian District, Beijing (39.95°N, 116.32°E) on May 27, 2015, and May 18, 2016 (Jiang et al., 2019). After emergence (VE), seedlings of similar size were selected so that each pot contained five uniform plants. Each variety was planted in three pots. The four developmental stages of soybean, namely, VE, R1, R7, and R8 (Fehr and Caviness, 1977), were investigated, and the average value of three pots for each variety was used for statistical analysis.

TABLE 1 | Soyabean varieties and maturity groups.

MGR (maturity group reference) varieties

Maturity group	North American varieties	Chinese-Russian varieties
MG 0000		Star4/75, Hujiao07-2479, Hujiao07-2123, Dongnong 36, Paula, R-4, Dongnong 41, Lingbei 8
MG 000	Maple Presto, OAC Vision, Rassvet, Jug 30, Mageva	R-2, Heihe 35
MG 00	Canatto, Maple Ridge, Daksoy, McCall, Agassiz	Mengdou 32, Beidou 16, Dongnong 44, Mengdou 11
MG 0	Traill, Chico, Barnes, Norpro, Dawson	Jiangmodou 1, Heihe 18, Heihe 43, Heihe 27, Beidou 37, Dengke 1, Fengshou 12, Dongnong 4, Hefeng 25
MG I	Haroson, Kato, Parker, Granite, NE1900	Heinong 16, Taixingheidou, Heinong 26, Suinong 14
MG II	Holt, Olympus, Century 84, IL1, LN92-7369	Jilin 20, Yongchengzihuadou, Xiangchundou 24, Tiefeng 19
MG III	Athow, LN89-5699, KS3494, IL2, Williams 82	Zhonghuang 13, Zhonghuang 30, Tiefeng 33, Zhongdou 39, Xudou 9, Tiefeng 31, Jindou 19, Huachun 6, Huaidou 9
MG IV	Flyer, Omaha, Calhoun, CF461, UA-4805	Zheng 92116, Guandou 2, Jindou 39, Shanning 16, Houzimao
MG V	Nathan, Hollady, Hutcheson, R01-3474F, TN04-5321	Shangdou 14, Dian 86-4, Diandou 7
MG VI	Desha, Musen, D95-6271, G01-PR16, Boggs	Zhongdou 38, Wuhuasiyuehuang, Suxiandou 19, Nannong 493-1
MG VII	Stonewall, Benning, Santee, Hagood	Huangfengwo, Tongshanbopihuang, Hengyangbayueqing, Nanxiadou 25
MG VIII	Motte, Dowling, Crockett, Prichard, IAC-8, CIGRAS-51, CIGRAS-06	Aijiaoqing, Pingguohuangdou, Nandou 12, Shangraodaqingsi, Lanxidaqingdou, Qiudou 1, Nandou 17, Jiangledaqingdou, Guixia 3
MG IX	Jupiter, Alamo, FT-15, UFV-3	
MG X	I.C. 192	Zigongdongdou

DNA Isolation, PCR, and Sequencing

Genomic DNA was extracted from fresh trifoliolate leaves using the standard cetyltrimethylammonium bromide (CTAB) method (Jiang et al., 2019). To amplify the 21 soybean *GmCOL* genes, 36 PCR primer pairs were used. *GmCOL9*/*15*/*16*/*20*/*28* was amplified with 2 primer pairs: *GmCOL4* with 11 primer pairs and *GmCOL1*/*2*/*3*/*5*/*6*/*8*/*10*/*13*/*14*/*19*/*22*/*23*/*24*/*25*/*26* with 1 primer pair. The sequences of these primers and the template designation are listed in **Supplementary Table 1**. Target regions were amplified with the high-fidelity polymerase of KOD-Plus-Neo and KOD-FX. Their reaction conditions were 94°C for 2 min (98°C for 10 s, 68°C for 3 min), 35 cycles, and 68°C for 10 min. The PCR products were directly sequenced using the Sanger method (TSINGKE Biological Technology Company, BGI and Omega Genetics, Beijing, China). The polymorphic site information of the 21 soybean *GmCOL* family genes is listed in **Supplementary Table 2**.

Data Mining and Sequence Analysis

The annotated soybean *COL* gene sequence was downloaded from the Joint Genome Institute (JGI) Phytozome website[1]. The protein sequences of the annotated *Arabidopsis* genes (TAIR9 release) were downloaded from the Arabidopsis Information Resource (TAIR) website[2]. The *CO*-like gene sequence data were also collected from the USDA-ARS Soybean Genetics and Genomics Database (SoyBase Database[3]). The reference sequence of the soybean variety Williams 82 was obtained from the Phytozome (version 12.0) database[4] as a reference design for the amplification and sequencing primers of the *COL* genes.

Sequencing was performed using an ABI3730 sequencer. Multiple sequence alignment, editing, and stitching were carried out using ClustalW in MEGA5 with default parameters. Single-nucleotide polymorphism (SNP) analysis was carried out using TASSEL5 software. The haplotype analysis of the sequencing results was conducted by DNAsp to determine whether the protein coding was affected. LD was also evaluated using TASSEL5 software. The statistical analysis was performed using "R" software.

RESULTS

Diversity of Soybean Varieties in Flowering Time and Maturity

The results of the growth traits are presented in **Table 2**. The varieties utilized in this experiment exhibited significant diversity in flowering time and maturity (**Table 2**). In 2015, 19 varieties, namely, MC119, MC63, MC69, MC70, MC54, MC121, MC122, MC123, MC72, MC124, MC125, MC126, MC48, MC64, MC65, MC66, MC67, MC68, and MC71, flowered but failed to reach R7. In 2016, 50 varieties did not reach R7 (**Table 2**). In 2015, the days to R1 from emergence (VE) ranged from 21 (MC82, MG 0) to 121 days (MC72, MG VIII) with a span of 100 days. In 2016, the span time was 96 days between MG 0000 (MC05, 13 days) and MG VIII (MC129, 109 days). In 2016, the days to R7 ranged from 68 (MC05 and MC74, MG 0000) to 144 days (MC42, MG IV), and the days to R8 ranged from 76 (MC05, MG 0000) to 109 days (MC37, MG III). The combination of maturity group information for the days of flowering time from the emergence and the days of beginning maturity from emergence in 2016 is plotted in a box diagram (**Figure 1**).

GmCOL Gene Sequencing and Mutation Study

A large number of mutated sites were found in *GmCOL2*, *GmCOL5*, *GmCOL14*, *GmCOL16*, *GmCOL20,* and *GmCOL28* (**Figures 2B1,D1** and **Supplementary Figures 1G1,I1, 2B1,H1**). A majority of the mutation sites were in gene intron regions, but a few were located in protein-coding regions caused by insertions-deletions (InDels) and substitutions. In addition, *GmCOL6*, *GmCOL22*, *GmCOL23*, *GmCOL24*, *GmCOL25,* and *GmCOL26* showed high mutation frequencies in the protein-coding regions (**Supplementary Figures 1, 2**). In contrast, *GmCOL4*, *GmCOL15,* and *GmCOL19* expressed lower mutational occurrences (**Supplementary Figures 1A1,H1, 2A1**). The exception was *GmCOL3*, which did not show mutations (**Figure 2C**) in either the intron or exon region of the entire gene. *GmCOL3* is a highly conserved gene with no change in sequence. However, the amino acid sequences encoded by *GmCOL1*, *GmCOL8*, *GmCOL9*, *GmCOL10,* and *GmCOL13* were not affected (**Figure 2A** and **Supplementary Figures 1C1-F1**).

Soybean *COL* Gene Family Exhibited Different Linkage Disequilibrium

Linkage and association mapping were drawn among the SNPs of the gene *via* TASSEL5. *GmCOL5* across the region almost from the starting to the end site presented a strong LD by completing the haplotype block (**Figure 3D**). *GmCOL20*, *GmCOL24,* and *GmCOL25* across the region expressed strong LD as a haplotype block (**Figures 3N,Q,R**), and the polymorphic sites of *GmCOL6*, *GmCOL8*, *GmCOL10*, *GmCOL13,* and *GmCOL19* showed a strong LD, with almost the entire region being a haplotype block (**Figures 3E,F,H,I,M**). Although *GmCOL2*, *GmCOL16*, *GmCOL22*, *GmCOL26,* and *GmCOL28* had a high-sequence polymorphism, the entire region distributed a few LDs (**Figures 3B,L,O,S,T**) with the same high-sequence polymorphism of *GmCOL4*, *GmCOL9,* and *GmCOL23*, and quite a few strong LDs were distributed broadly (**Figures 3C,G,P**). *GmCOL1*, *GmCOL14,* and *GmCOL15* did not reveal any LD (**Figures 3A,J,K**). *GmCOL3* was the most conserved gene, showing no polymorphic site(s).

Haplotype Analysis and Its Association With Maturity Groups of Soybean Varieties

Based on LD and sequence clustering, haplotype definition was analyzed for the 21 soybean *GmCOL* gene families in

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

[2]www.arabidopsis.org

[3]http://soybase.org/

[4]http://phytozome.jgi.doe.gov/pz/portal.html

TABLE 2 | Phenotypic variation of soybean varieties at Beijing.

Code	Variety name	MG	Phenotype in 2016		
			VE-R1	VE-R7	VE-R8
MC01	Star 4/75	0000	15.4 ± 1.2	70.6 ± 1.3	80.6 ± 0.61
MC02	Hujiao 07-2479	0000	16.4 ± 0.82	72.06 ± 1.27	81.26 ± 0.45
MC03	Hujiao 07-2123	0000	15.6 ± 0.50	71.13 ± 0.35	80.8 ± 0.41
MC04	Dongnong 36	0000	14.13 ± 0.51	69.13 ± 0.51	79.66 ± 0.97
MC05	Paula	0000	13.66 ± 0.48	68.4 ± 0.82	76.66 ± 0.48
MC06	R-4	0000	25.86 ± 0.74	71.46 ± 0.91	82.6 ± 0.82
MC07	Dongnong 41	0000	14.33 ± 0.48	69.46 ± 0.51	79.53 ± 0.91
MC74	Lingbei 8	0000	18.9 ± 1.0	68.4 ± 1.2	80.2 ± 1.5
MC75	Dongnong 41-C	0000	25.7 ± 1.4	70.5 ± 0.9	80.9 ± 1.0
MC08	Maple-Presto	000	15.53 ± 0.51	69.2 ± 1.01	77.66 ± 0.97
MC09	OAC-Vision	000	16.73 ± 0.79	69.8 ± 1.52	77.53 ± 0.91
MC10	Rassvet	000	17.33 ± 0.48	71.53 ± 0.91	79.8 ± 0.41
MC11	Jug-30	000	24.33 ± 0.48	71.2 ± 0.41	79.8 ± 0.67
MC12	Mageva	000	25.73 ± 0.45	75.66 ± 0.97	83.8 ± 0.86
MC73	R2	000	14.6 ± 0.9	70.6 ± 1.4	80.8 ± 2.1
MC76	Heihe 35	000	25.4 ± 1.5	73.5 ± o.5	87.5 ± 0.5
MC13	Canatto	00	18.2 ± 0.4	72.8 ± 1.8	85.2 ± 0.4
MC14	Maple-Ridge	00	18.4 ± 0.5	72.0 ± 1.0	84.7 ± 0.4
MC15	Daksoy	00	24.3 ± 0.4	71.3 ± 0.4	80.8 ± 0.3
MC16	McCall	00	23.6 ± 0.8	73.4 ± 1.2	85.8 ± 1.5
MC17	Agassiz	00	25.2 ± 0.4	71.6 ± 0.8	81.2 ± 0.4
MC77	Mengdou 32	00	22.6 ± 2.4	73.2 ± 1.0	86.3 ± 0.9
MC78	Beidou 16	00	19.9 ± 0.7	73.4 ± 1.5	85.4 ± 0.5
MC80	Dongnong 44	00	24.8 ± 1.0	71.3 ± 1.4	81.3 ± 0.9
MC81	Mengdou 11	00	24.3 ± 0.4	71.2 ± 1.3	81.6 ± 1.5
MC26	Traill	0	27.3 ± 0.4	73.3 ± 0.9	82.26 ± 0.7
MC27	Chico	0	26.4 ± 0.9	70.8 ± 1.5	81.4 ± 1.5
MC28	Barnes	0	26.5 ± 0.7	68.8 ± 1.3	78.4 ± 0.5
MC29	Norpro	0	25.3 ± 0.9	71.4 ± 0.50	82.1 ± 1.2
MC30	Dawson	0	23.6 ± 0.4	70.2 ± 1.37	80.4 ± 1.2
MC31	Jiangmodou 1	0	23.4 ± 0.9	69.7 ± 1.1	80.5 ± 0.9
MC82	Heihe 18	0	25.2 ± 1.5	72.0 ± 1.0	80.8 ± 1.3
MC83	Heihe 43	0	26.4 ± 0.5	71.9 ± 1.6	82.6 ± 0.7
MC84	Heihe 27	0	23.6 ± 0.4	70.8 ± 1.0	81.6 ± 0.5
MC85	Beidou 37	0	21.5 ± 0.9	69.5 ± 0.9	80.2 ± 1.2
MC86	Dengke1	0	23.8 ± 0.4	69.6 ± 0.4	80.4 ± 0.5
MC97	Fengshou 12	0	24.6 ± 0.9	71.0 ± 1.0	80.8 ± 1.0
MC88	Dongnong-4	0	32.2 ± 1.0	74.1 ± 0.8	85.8 ± 0.6
MC89	Hefeng 25	0	32.7 ± 0.4	75.5 ± 0.9	85.7 ± 0.7
MC23	Haroson	I	25.2 ± 0.45	71.2 ± 0.4	82.2 ± 0.4
MC24	Kato	I	27.5 ± 0.9	72.2 ± 0.4	81.5 ± 0.9
MC25	Parker	I	25.4 ± 0.9	71.4 ± 0.9	80.3 ± 0.9
MC26	Granite	I	25.3 ± 0.9	71.3 ± 0.9	78.4 ± 0.9
MC27	NE1900	I	35.2 ± 0.4	70.7 ± 1.0	80.1 ± 0.3
MC90	Heinong 16	I	28.5 ± 0.9	71.6 ± 0.9	79.4 ± 0.5
MC91	Taixingheidou	I	26.6 ± 0.9	72.73 ± 1.6	79.3 ± 1.7
MC92	Heinong 26	I	29.4 ± 0.5	70.8 ± 0.6	78.8 ± 0.9
MC93	Suinong 14	I	27.2 ± 1.5	69.4 ± 1.3	78.2 ± 1.3
MC28	Holt	II	30.3 ± 2.4	70.8 ± 0.8	81.8 ± 1.5
MC29	Olympus	II	27.7 ± 0.9	72.3 ± 0.4	82.0 ± 1.0
MC30	Century-84	II	26.6 ± 0.4	71.6 ± 0.8	81.0 ± 1.2
MC31	IL1	II	30.4 ± 5.5	70.2 ± 1.7	79.6 ± 2.5
MC32	LN92-7369	II	30.4 ± 1.4	76.2 ± 1.0	107.9 ± 1.6
MC94	Jilin 20	II	25.2 ± 1.0	70.3 ± 0.9	78.2 ± 1.0
MC95	Yongchengzihuadou	II	41.3 ± 1.2	72.1 ± 1.2	87.7 ± 1.2
MC96	Xiangchundou 24	II	43.8 ± 1.0	77.4 ± 0.5	86.6 ± 1.2
MC98	Tiefeng 19	II	24.4 ± 0.5	71.2 ± 1.2	81.2 ± 1.7
MC33	Athow	III	30.3 ± 0.9	76.3 ± 0.9	106.3 ± 0.9
MC34	Zhonghuang 13	III	43.3 ± 0.4	78.4 ± 0.9	106.4 ± 0.9
MC35	Zhonghuang 30	III	37.8 ± 0.9	78.3 ± 0.9	105.5 ± 1.4
MC36	LN89-5699	III	32.3 ± 0.9	80.2 ± 1.5	108.2 ± 0.4
MC37	KS3494	III	37.4 ± 0.5	79.8 ± 0.9	109.4 ± 1.5
MC38	IL2	III	42.1 ± 0.8	105.2 ± 2.3	105.3 ± 5.1
MC39	Williams 82	III	36.8 ± 0.3	107.3 ± 0.8	108.1 ± 2.3

(Continued)

TABLE 2 | (Continued)

Code	Variety name	MG	Phenotype in 2016		
			VE-R1	VE-R7	VE-R8
MC97	Tiefeng 33	III	29.5 ± 0.5	78.7 ± 0.4	106.8 ± 1.2
MC99	Zhongdou 39	III	44.5 ± 0.5	77.4 ± 0.5	105.2 ± 1.8
MC100	Xudou 9	III	37.9 ± 0.7	78.0 ± 1.0	106.4 ± 1.5
MC101	Tiefeng 31	III	24.2 ± 1.1	77.2 ± 1.3	105.2 ± 1.1
MC102	Jindou 19	III	25.8 ± 1.3	77.4 ± 1.5	104.8 ± 1.5
MC103	Huachun 6	III	31.4 ± 0.5	77.3 ± 0.4	105.6 ± 0.9
MC104	Huaidou 9	III	34.3 ± 0.4	ND	ND
MC40	Flyer	IV	35.5 ± 1.4	143.0 ± 1.4	ND
MC41	Omaha	IV	38.0 ± 1.0	143.4 ± 1.4	ND
MC42	Calhoun	IV	37.3 ± 0.4	144.4 ± 0.5	ND
MC43	CF461	IV	31.8 ± 1.0	143.7 ± 0.4	ND
MC44	UA-4805	IV	61.2 ± 0.4	ND	ND
MC105	Zheng 92116	IV	37.2 ± 0.4	ND	ND
MC106	Guandou 2	IV	35.8 ± 1.0	ND	ND
MC107	Jindou 39	IV	36.2 ± 1.3	142.9 ± 1.0	ND
MC108	Shanning 16	IV	43.4 ± 0.8	142.3 ± 0.4	ND
MC109	Houzimao	IV	43.7 ± 1.4	143.3 ± 1.1	ND
MC45	Nathan	V	60.2 ± 1.0	ND	ND
MC46	Holladay	V	62.2 ± 1.5	ND	ND
MC47	Hutcheson	V	63.2 ± 1.0	ND	ND
MC48	R01-3474F	V	62.2 ± 1.5	ND	ND
MC49	TN04-5321	V	63.6 ± 1.5	ND	ND
MC110	Shangdou 14	V	44.7 ± 0.4	ND	ND
MC111	Dian 86-4	V	43.6 ± 1.5	ND	ND
MC113	Diandou 7	V	88.3 ± 0.9	ND	ND
MC50	Desha	VI	48.0 ± 0.8	ND	ND
MC51	Musen	VI	80.0 ± 1.0	ND	ND
MC52	D95-6271	VI	82.3 ± 1.5	ND	ND
MC53	G01-PR16	VI	81.6 ± 0.9	ND	ND
MC54	Boggs	VI	79.8 ± 1.5	ND	ND
MC112	Zhongdou 38	VI	83.6 ± 2.5	ND	ND
MC114	Wuhuasiyuehuang	VI	86.3 ± 0.9	ND	ND
MC108	Suxiandou 19	VI	85.8 ± 1.2	ND	ND
MC116	Nannong 493-1	VI	91.2 ± 0.4	ND	ND
MC55	Stonewall	VII	81.4 ± 0.5	ND	ND
MC57	Santee	VII	82.6 ± 1.8	ND	ND
MC58	Hagood	VII	94.6 ± 1.2	ND	ND
MC56	Benning	VII	93.3 ± 0.9	ND	ND
MC117	Hunagfengwu	VII	93.7 ± 0.4	ND	ND
MC118	Tongshanbopihuang	VII	93.2 ± 1.2	ND	ND
MC60	Motte	VIII	93.3 ± 0.9	ND	ND
MC61	Dowling	VIII	95.8 ± 1.5	ND	ND
MC62	Crockett	VIII	65.4 ± 0.2	142.2 ± 0.5	158.0 ± 1.0
MC128	Hengyangbayueqing	VII	96.9 ± 1.0	ND	ND
MC119	Nanxiadou 25	VII	92.6 ± 0.9	ND	ND
MC63	Prichard	VIII	94.8 ± 1.0	ND	ND
MC69	CIGRAS-51	VIII	81.4 ± 0.5	ND	ND
MC70	CIGRAS-06	VIII	81.5 ± 0.9	ND	ND
MC54	Aijiaoqing	VIII	94.3 ± 0.4	ND	ND
MC121	Pingguohuangdou	VIII	92.9 ± 1.0	ND	ND
MC122	Nandou 12	VIII	91.3 ± 0.4	ND	ND
MC123	Shangraodaqingsi	VIII	92.8 ± 1.0	ND	ND
MC72	Zigongdongdou	VIII	73.4 ± 1.5	ND	ND
MC124	Lanxidaqingdou	VIII	91.6 ± 0.5	ND	ND
MC125	Qiudou 1	VIII	94.3 ± 0.4	ND	ND
MC126	Nandou 17	VIII	91.6 ± 0.4	ND	ND
MC48	Jiangledaqingdou	VIII	99.5 ± 0.9	ND	ND
MC129	Guixia 3	VIII	109.2 ± 1.0	118.4 ± 0.2	ND
MC64	Jupiter	IX	96.3 ± 0.9	ND	ND
MC65	Alamo	IX	96.0 ± 1.3	ND	ND
MC66	FT-15	IX	94.2 ± 1.3	ND	ND
MC67	UFV-3	IX	96.2 ± 1.0	ND	ND
MC68	IAC-8	IX	90.0 ± 1.0	ND	ND
MC71	I.C.-192	X	80.4 ± 1.5	ND	ND

ND, Not available data; VE, Variation after emergence.

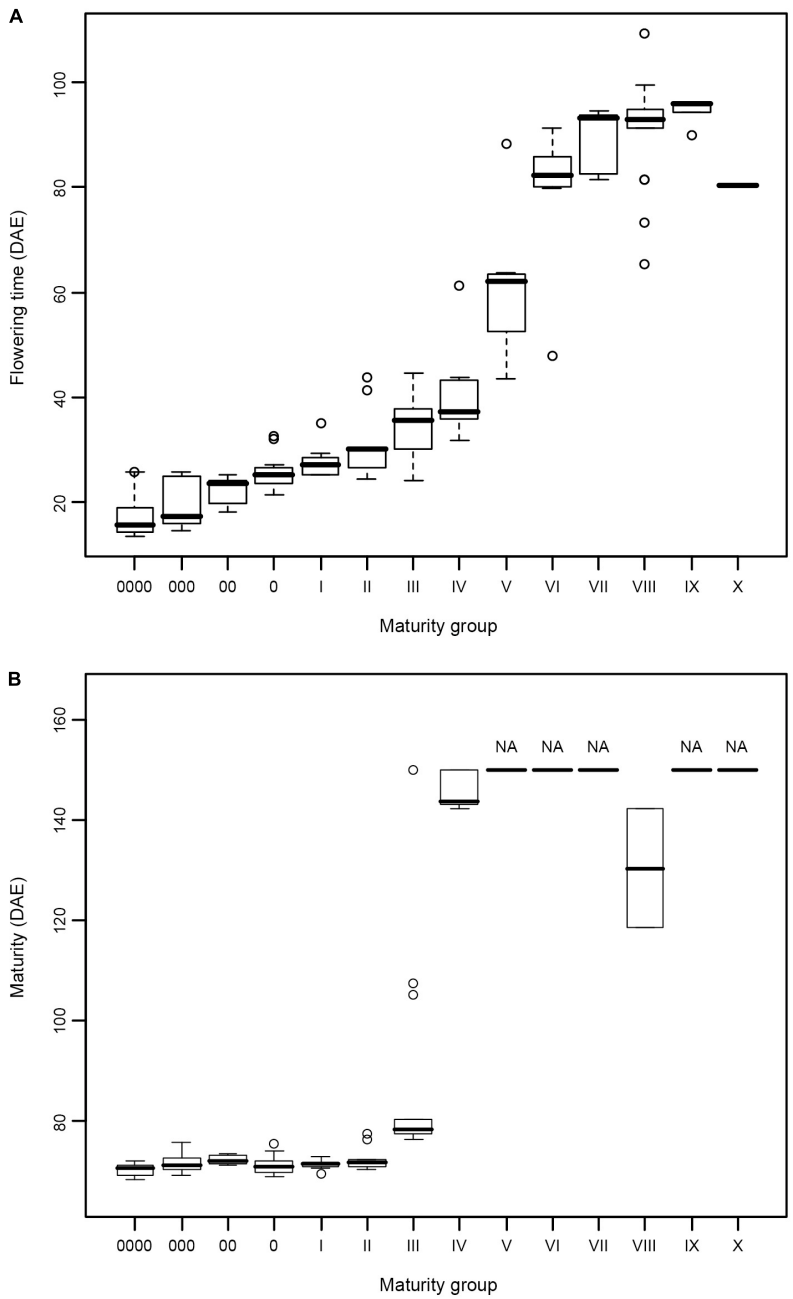


FIGURE 1 | (A,B) Flowering time and maturity of soybean varieties from different maturity groups. Flowering time and maturity data were collected in 2016 from 128 soybean varieties covering 14 maturity groups. DAE, days after emergence.

the 128 varieties investigated. A total of 21 haplotypes of the *GmCOL* gene family variation types and polymorphic sites used for composing haplotypes are listed in **Figure 2** and **Supplementary Tables 3, 4** (**Supplementary Figures 1, 2**), respectively. **Supplementary Table 5** shows 21 *GmCOL* protein groups associated with haplotypes.

GmCOL2-Hap2 was dominant in the 41 accessions covering all maturity groups, except MG X. *GmCOL2-Hap4* showed early flowering compared with the other haplotype groups (**Figures 2B3,B4**, **Table 2**, and **Supplementary Table 3**). *GmCOL4-Hap1* accounted for 99 varieties, with all maturity groups excluding MG X (**Supplementary Figure 1A2**). *GmCOL4-Hap1* had an earlier flowering date in 2016. *GmCOL5-Hap2* was present in 56 accessions, covering all maturity groups (**Figure 2D1**). *GmCOL6-Hap1* was the most abundant in 70 varieties from different maturity groups, including MG 0000-MG V and MG VIII-MG X (**Supplementary Figure 1B1**). Among the 11 haplotypes, *GmCOL9-Hap1* was

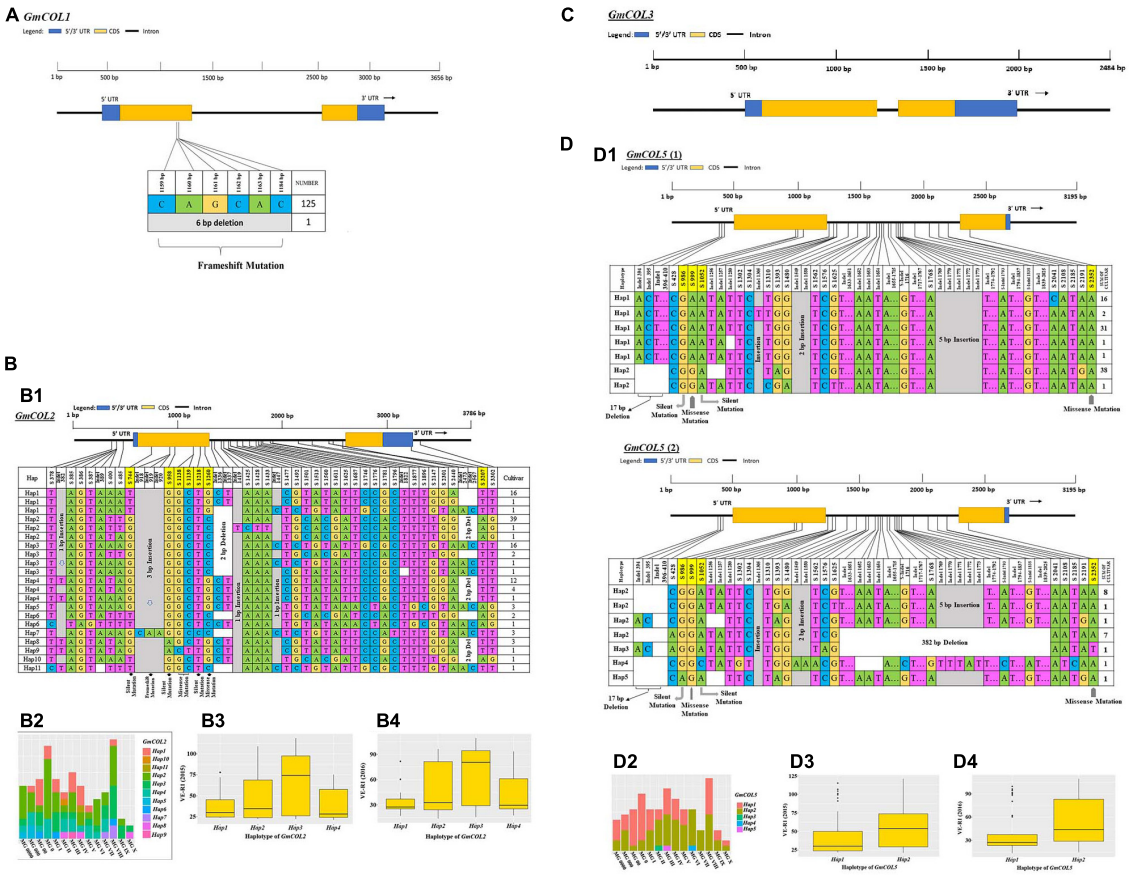


FIGURE 2 | Haplotype analysis, distribution in different maturity groups, and flowering time (VE-R1) in the main haplotypes of approximately four *GmCOL* family genes. **(A)** Haplotype of *GmCOL1*. **(B,B1)** Haplotypes of *GmCOL2*. The haplotype is shown as a linear combination of alleles. The site combination labeled with the yellow outline represents the tagging haplotype (the composition of the haplotype depends on fewer and critical variation sites). **(B2)** Haplotype distribution of *GmCOL2* in different maturity groups. **(B3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL2* in 2015. **(B4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL2* in 2016. **(C)** Gene structure of *GmCOL3*. **(D,D1)** Haplotypes of *GmCOL5* (1) and *GmCOL5* (2). **(D2)** Haplotype distribution of *GmCOL5* in different maturity groups. **(D3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL5* in 2015. **(D4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL5* in 2016.

dominant in 67 varieties from maturity groups MG 0000-MG X (**Supplementary Figures 1D1,D2**). *GmCOL9-Hap3* and *GmCOL9-Hap4* were distributed in a few varieties, and the flowering time for both haplotypes was early. *GmCOL10-Hap1* was the most abundantly expressed in 91 varieties from all maturity groups except MG X (**Supplementary Figures 1E1,E2**). *GmCOL16-Hap1* was rich in accessions distributed in the maturity groups MG 0000-MG VIII (**Supplementary Figures 1I1,I2**). *GmCOL20-Hap2* was present in 52 varieties from all maturity groups (MG 0000-MG X) (**Supplementary Figures 2B1,B2**). Among all the haplotypes of *GmCOL28*, *GmCOL28-Hap1* was the most common, accounting for 59 accessions, and was distributed in the maturity groups MG 0000, MG 00, MG 0, MG I-MG VIII, and MG X (**Supplementary Figures 2H1,H2**). In this experiment, no haplotype variants were found in *GmCOL1* and *GmCOL3* (**Figures 2A,C**). The *Hap1* series of haplotypes of *GmCOL8/13/14/15/19/22/23/24/25/26* was the most abundant in the varieties and was distributed in all 14 maturity groups (**Supplementary Figures 1C1,F1,G1,H1, 2A1,C1,D1,E1,F1,G1**).

Haplotypes Associated With Flowering Time

The haplotype groups of 21 gene families are presented in **Table 3**. An analysis of variance (ANOVA) was performed to elucidate the natural variations in flowering time (VE-R1) of the *GmCOL* gene families based on haplotype groups. In this analysis, eight genes *(GmCOL2/5/9/13/15/16/25/28)* showed significant results from the first emergence to the first flowering date (VE-R1) in both years, and three genes [*GmCOL8* (2015), *GmCOL22* (2015), and *GmCOL26* (2016)] exhibited significant results in a single year (**Supplementary Table 6**). The remaining genes (*GmCOL4/6/10/14/19/20/23/24*) did not show significant results in these years. Notably, *GmCOL1* and *GmCOL3* were the most conserved, and no polymorphisms in the coding region were observed.

Among the 11 haplotype groups of *GmCOL2*, 4 were analyzed, and *GmCOL2-Hap3* was significantly different from every other haplotype group (**Figure 4A**). The *GmCOL2-Hap1/2/4* haplotype groups, which appeared in the above haplotype group, were not significantly different from each other. Analysis of *GmCOL5*

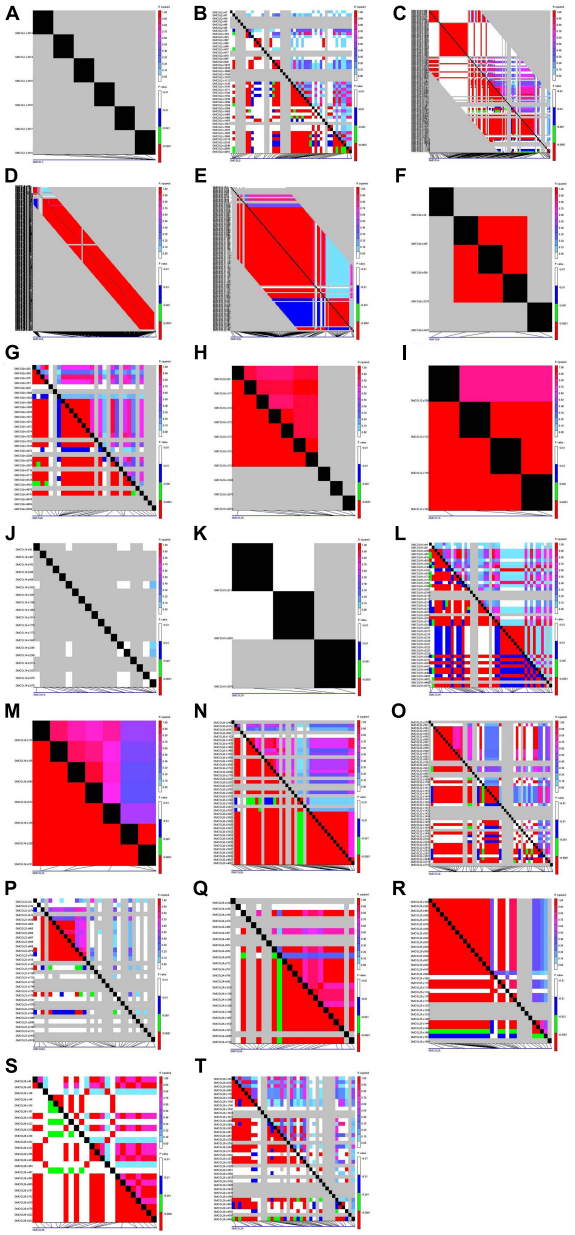


FIGURE 3 | Linkage disequilibrium of *GmCOL* family genes in soybean. **(A)** Linkage disequilibrium (LD) of *GmCOL1*. **(B)** LD of *GmCOL2*. **(C)** LD of *GmCOL4*. **(D)** LD of *GmCOL5*. **(E)** LD of *GmCOL6*. **(F)** LD of *GmCOL8*. **(G)** LD of *GmCOL9*. **(H)** LD of *GmCOL10*. **(I)** LD of *GmCOL13*. **(J)** LD of *GmCOL14*. **(K)** LD of *GmCOL15*. **(L)** LD of *GmCOL16*. **(M)** LD of *GmCOL19*. **(N)** LD of *GmCOL20*. **(O)** LD of *GmCOL22*. **(P)** LD of *GmCOL23*. **(Q)** LD of *GmCOL24*. **(R)** LD of *GmCOL25*. **(S)** LD of *GmCOL26*. **(T)** LD of *GmCOL28*.

showed that in both years two major haplotypes showed significant differences from each other, and *Hap2* was related to late flowering compared with *Hap1* (**Figure 4B**). In terms of *GmCOL5 Hap2* (S999), the original nucleotide (A^{499}) was mutated to G^{499} in the coding sequence, resulting in a missense mutation in the amino acid sequence (K^{167} to E^{167}). K^{167} was located between the B-box and the CCT domain. In comparison, *GmCOL9*, *Hap1,* and *Hap2* (**Figure 4C**) did not show significant differences from each other, and both were significantly different from *GmCOL9-Hap3* and *GmCOL9-Hap4* only in 2015. Likewise, *GmCOL13-Hap2/3* (**Figure 4D**) was closely associated with flowering time because both were significantly different from *GmCOL13-Hap1* in both years.

For *GmCOL13 Hap2/3* (S584), the original nucleotide (G^{84}) was mutated to A^{84} in the coding sequence, resulting in a silent mutation in the no. 28 amino acid sequence, which was located in the first B-box. In both years, *GmCOL15-Hap3/4* (**Figure 4E**) showed a significant difference compared with *GmCOL15-Hap1/2*. For *GmCOL16*, the haplotypes showed non-significant results among the four haplotypes (*Hap1*, *Hap2*, *Hap3,* and *Hap4*) (**Figure 4F**). In the first year, five haplotypes of *GmCOL25* (*Hap1*, *Hap2*, *Hap3*, *Hap4,* and *Hap5*) were not significantly different from each other, whereas *Hap3* and *Hap4* were significantly different from *Hap1*, *Hap2,* and *Hap5* in the second year (**Figure 4G**). *Hap2* of *GmCOL28* was significantly different from *Hap1* and *Hap3* in both years (**Figure 4H**). For *GmCOL28 Hap2* (S1066), the original nucleotide (T^{263}) was mutated to C^{263} in the coding sequence, resulting in a missense mutation in the amino acid sequence (V^{88} to A^{88}), which was located in the first B-box, and may affect the protein and protein interaction. Thus, the results revealed that *Hap2 of GmCOL28* was closely associated with flowering time.

DISCUSSION

Selected Soybean Varieties Showed Diversity in Flowering Time and Maturity

The varieties analyzed in this study showed the diversity of flowering time and maturity, which indicated that the flowering time (VE-R1) and the days from emergence to physiological maturity (VE-R7) were related to the variety trait and natural environment (Jiang et al., 2019). According to the photoperiod responses, many soybean varieties have evolved to adapt to a broad range of growing areas in China (Wang et al., 2016). The flowering time of the selected varieties ranged from 21 to 121 days and 13.66 to 109.2 days, and the maturity time ranged from 61.2 to 150.7 days and 68.4 to 144.4 days, respectively, for two consecutive years (2015 and 2016) (Jiang et al., 2019), indicating high diversity (**Table 2**). However, Paula showed the shortest VE-R1 in the 2nd year, with the maturity group MG 0000. Paula is recognized as a high-latitude cold region (HCR) soybean variety and HCR soybean, which typically matures early (Jia et al., 2014). This observation suggested that the maturity group MG 0000 was relatively stable for days to start flowering from emergence.

GmCOL Orthologs Showed Divergence in Sequence Polymorphism

It has been shown that GmCOL1a (Glyma08g28370) and GmCOL1b (Glyma18g51320) in soybean are the closest *Arabidopsis* COL2 orthologs to CO (Thakare et al., 2010). The

TABLE 3 | Haplotype frequency of *GmCOL* family genes.

	Haplotype										
GmCOL2	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*	*Hap7*	*Hap8*	*Hap9*	*Hap10*	*Hap11*
Accessions	18	41	20	17	3	3	3	3	1	1	1
GmCOL4	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	99	7	4	4	1						
GmCOL5	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	51	56	1	1	1						
GmCOL6	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	70	46	5	1	1						
GmCOL8	*Hap1*	*Hap2*									
Accessions	88	11									
GmCOL9	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*	*Hap7*	*Hap8*	*Hap9*	*Hap10*	*Hap11*
Accessions	67	25	4	4	3	1	1	1	1	1	1
GmCOL10	*Hap1*	*Hap2*	*Hap3*	*Hap4*							
Accessions	91	16	1	1							
GmCOL13	*Hap1*	*Hap2*	*Hap3*								
Accessions	122	5	1								
GmCOL14	*Hap1*	*Hap2*	*Hap3*								
Accessions	114	1	1								
GmCOL15	*Hap1*	*Hap2*	*Hap3*	*Hap4*							
Accessions	107	5	2	1							
GmCOL16	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*	*Hap7*	*Hap8*	*Hap9*		
Accessions	54	25	24	6	5	4	2	2	1		
GmCOL19	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	99	17	2	2	1						
GmCOL20	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*	*Hap7*	*Hap8*	*Hap9*	*Hap10*	
Accessions	19	52	27	8	2	2	1	1	1	1	
GmCOL22	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*					
Accessions	110	7	4	3	2	1					
GmCOL23	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*	*Hap6*	*Hap7*	*Hap8*	*Hap9*	*Hap10*	
Accessions	80	4	3	2	1	1	1	1	1	1	
GmCOL24	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	88	5	3	1	1						
GmCOL25	*Hap1*	*Hap2*	*Hap3*	*Hap4*	*Hap5*						
Accessions	89	9	6	6	1						
GmCOL26	*Hap1*	*Hap2*	*Hap3*	*Hap4*							
Accessions	61	56	7	2							
GmCOL28	*Hap1*	*Hap2*	*Hap3*								
Accessions	59	29	22								

transgenic soybean line overexpressing *GmCOL1a* flowered late under long-day or natural conditions (Cao et al., 2015). There is a single-nucleotide substitution at the GmCOL1b CCT domain in the *gmcol1b* mutant, leading to the mutagenesis of conserved threonine at amino acid 314 into isoleucine, which results in early flowering of the *gmcol1b* mutant (Cao et al., 2015). In this study, the varieties that were taken for the observation of natural variations regarding flowering time and maturation showed divergent characteristics in flowering time (Jiang et al., 2019). Allelic variation can have an effect on flowering time (Irwin et al., 2016). In this study, the insertion and single-nucleotide substitution (InDel918, InDel919, InDel920, s958, s1138, s1139, s1218, and s1260) found in the first exon in *GmCOL2* led to missense mutations, which might have an effect on flowering time and maturity. The other genes in the *GmCOL* family showed various types of point mutations (**Figure 2** and **Supplementary Figures 1–5**). The loss of functions may have an impact on gene function, and the presence of these mutations suggested that polymorphism could be the main cause of the diversity of soybean flowering time.

Polymorphism of *GmCOL* Family Genes Was Associated With Flowering Time and Maturity

Haplotype-based analyses have been successfully carried out in different crops, such as maize (*Zea mays* L.) (Weber et al., 2009; Van Inghelandt et al., 2012; Lipka et al., 2013), rice (*Oryza sativa* L.) (Lestari et al., 2011; Yonemaru et al., 2012, 2014; Shao et al., 2013; Wu et al., 2020), and soybean (Choi et al., 2007; Li et al., 2009; Langewisch et al., 2014; Patil et al., 2016; Jiang et al., 2019). In this study, some polymorphic sites were identified and utilized by conducting a haplotype analysis of 21 soybean *GmCOL* gene families. Based on LD analysis and polymorphic sites, tagging haplotypes composed of some SNPs and InDels indicated that there were 7

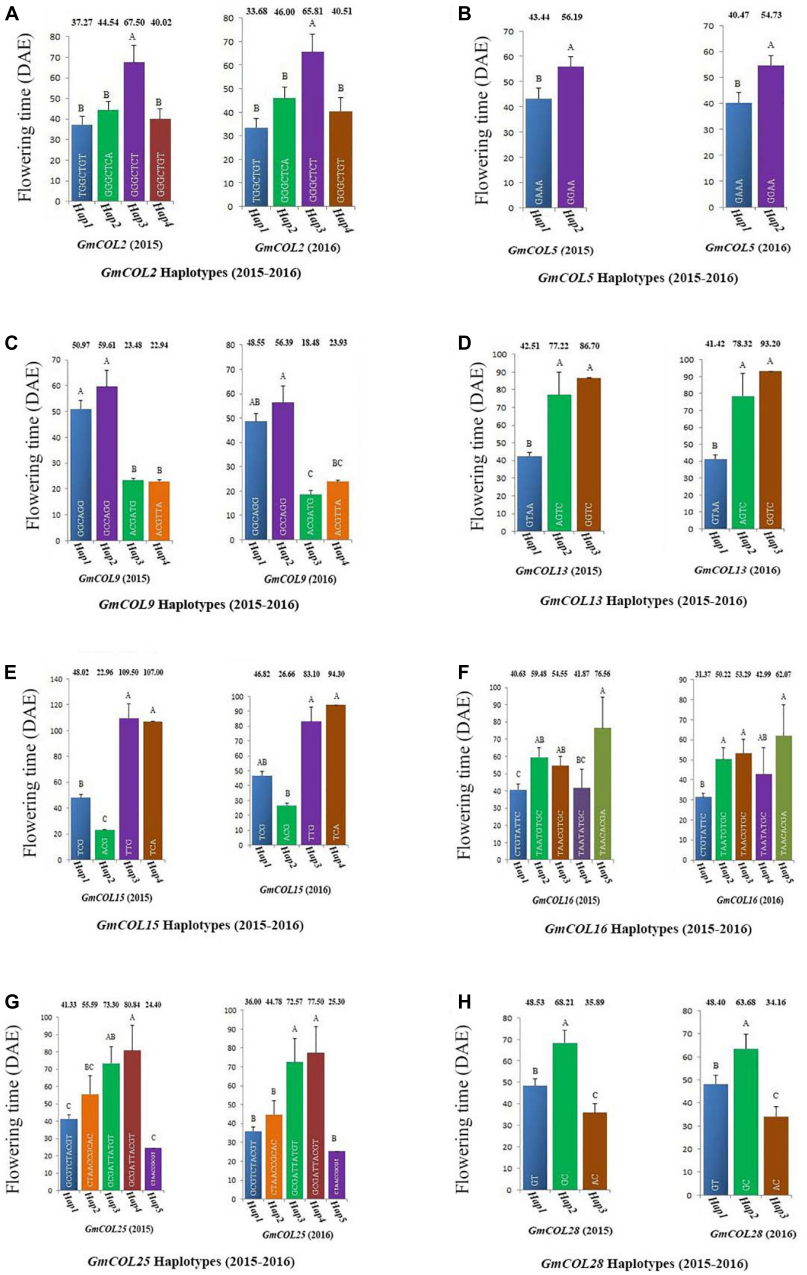


FIGURE 4 | Haplotypes of *GmCOL* genes in relation to flowering time (2015–2016). The relationship between flowering time and haplotype of *GmCOL2* **(A)**, *GmCOL5* **(B)**, *GmCOL9* **(C)**, *GmCOL13* **(D)**, *GmCOL15* **(E)**, *GmCOL16* **(F)**, *GmCOL25* **(G)**, and *GmCOL28* **(H)**.

polymorphic sites in *GmCOL2* and just 3 polymorphic sites in *GmCOL10* and 46 and 10 polymorphic sites in *GmCOL2* and *GmCOL10*, respectively (**Supplementary Table 4**). In addition, some haplotypes associated with flowering date and maturity (such as *GmCOL2-Hap1*, *Hap2*, and *Hap4*; *GmCOL9-Hap1*, *Hap2*, *Hap3*, and *Hap4*; *GmCOL13-Hap1*, *Hap2*, and *Hap3*; *GmCOL15-Hap1*, *Hap2*, *Hap3*, and *Hap4*; *GmCOL16-Hap1*, *Hap2*, *Hap3*, *Hap4*, and *Hap5*; and *GmCOL25-Hap1*, *Hap2*, *Hap3*, *Hap4*, and *Hap5*) were distributed in varieties with different maturity groups. The LD differences among the 21 soybean *COL* gene families may reflect selection pressure and, to some extent, the process of natural selection and domestication.

CONCLUSION

In summary, 21 *GmCOL* genes exhibited natural divergence in association with flowering and growth periods. Significant changes were found in 21 *COL* genome sequences, among which *GmCOL1*, *GmCOL3*, *GmCOL8*, *GmCOL9*, *GmCOL10,* and

GmCOL13 were conserved, but the sequences of the remaining *COL* genes showed a wide range of changes that could alter their function. In this study, it was noticed that additional polymorphisms linked to the 21 *GmCOL* gene families in trait-controlling regions might have a significant impact on flowering time and maturity.

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

MA conducted the major phenotypic and genotypic works. SZ, RE, WS, TW, and BJ joined the experiments. SY, CW, WH, SS, and YF provided the materials and technical supports. MA, RE, and BJ prepared the initial draft of the manuscript. FC rewrote the manuscript. FC, BJ, MA, RE, and TH revised the manuscript. TH coordinated and supervised the project. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Key R&D Program of China (2021YFF1001203 and 2017YFD0101400), China Agriculture Research System (CARS-04), and CAAS Agricultural Science and Technology Innovation Project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.817544/full#supplementary-material

Supplementary Figure 1 | Haplotype analysis, distribution in different maturity groups, and flowering time (VE-R1) in the main haplotypes of approximately nine *GmCOL* family genes. **(A,A1)** Haplotypes of *GmCOL4*. **(A2)** Haplotype distribution of *GmCOL4* in different maturity groups. **(A3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL4* in 2015. **(A4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL4* in 2016. **(B,B1)** Haplotypes of *GmCOL6*. **(B2)** Haplotype distribution of *GmCOL6* in different maturity groups. **(B3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL6* in 2015. **(B4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL6* in 2016. **(C,C1)** Haplotypes of *GmCOL8*. **(C2)** Haplotype distribution of *GmCOL8* in different maturity groups. **(C3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL8* in 2015. **(C4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL8* in 2016. **(D,D1)** Haplotypes of *GmCOL9* (1) and *GmCOL9* (2). **(D2)** Haplotype distribution of *GmCOL9* in different maturity groups. **(D3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL9* in 2015. **(D4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL9* in 2016. **(E,E1)** Haplotypes of *GmCOL10*. **(E2)** Haplotype distribution of *GmCOL10* in different maturity groups. **(E3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL10* in 2015. **(E4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL10* in 2016. **(F,F1)** Haplotypes of *GmCOL13*. **(F2)** Haplotype distribution of *GmCOL13* in different maturity groups. **(F3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL13* in 2015. **(F4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL13* in 2016. **(G,G1)** Haplotypes of *GmCOL14*. **(G2)** Haplotype distribution of *GmCOL14* in different maturity groups. **(G3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL14* in 2015. **(G4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL14* in 2016. **(H,H1)** Haplotypes of *GmCOL15*. **(H2)** Haplotype distribution of *GmCOL15* in different maturity groups. **(H3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL15* in 2015. **(H4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL15* in 2016. **(I,I1)** Haplotypes of *GmCOL16* (1), *GmCOL16* (2), *GmCOL16* (3), and *GmCOL16* (4). **(I2)** Haplotype distribution of *GmCOL16* in different maturity groups. **(I3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL16* in 2015. **(I4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL16* in 2016.

Supplementary Figure 2 | Haplotype analysis, distribution in different maturity groups, and flowering time (VE-R1) in the main haplotypes of approximately eight *GmCOL* family genes. **(A,A1)** Haplotypes of *GmCOL19*. **(A2)** Haplotype distribution of *GmCOL19* in different maturity groups. **(A3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL19* in 2015. **(A4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL19* in 2016. **(B,B1)** Haplotypes of *GmCOL20* (1), *GmCOL20* (2), and *GmCOL20* (3). **(B2)** Haplotype distribution of *GmCOL20* in different maturity groups. **(B3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL20* in 2015. **(B4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL20* in 2016. **(C,C1)** Haplotypes of *GmCOL22* (1) and *GmCOL22* (2). **(C2)** Haplotype distribution of *GmCOL22* in different maturity groups. **(C3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL22* in 2015. **(C4)** Flowering time (VER1) of the main haplotypes of *GmCOL22* in 2016. **(D,D1)** Haplotypes of *GmCOL23* (1) and *GmCOL23* (2). **(D2)** Haplotype distribution of *GmCOL23* in different maturity groups. **(D3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL23* in 2015. **(D4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL23* in 2016. **(E,E1)** Haplotypes of *GmCOL24*. **(E2)** Haplotype distribution of *GmCOL24* in different maturity groups. **(E3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL24* in 2015. **(E4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL24* in 2016. **(F,F1)** Haplotypes of *GmCOL25*. **(F2)** Haplotype distribution of *GmCOL25* in different maturity groups. **(F3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL25* in 2015. **(F4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL25* in 2016. **(G,G1)** Haplotypes of *GmCOL26*. **(G2)** Haplotype distribution of *GmCOL26* in different maturity groups. **(G3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL26* in 2015. **(G4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL26* in 2016. **(H,H1)** Haplotypes of *GmCOL28* (1) and *GmCOL28* (2). **(H2)** Haplotype distribution of *GmCOL28* in different maturity groups. **(H3)** Flowering time (VE-R1) of the main haplotypes of *GmCOL28* in 2015. **(H4)** Flowering time (VE-R1) of the main haplotypes of *GmCOL28* in 2016.

REFERENCES

Alonso-Blanco, C., Mendez-Vigo, B., and Koornneef, M. (2005). From phenotypic to molecular polymorphisms involved in naturally occurring variation of plant development. *Int. J. Dev. Biol.* 49, 717–732. doi: 10.1387/ijdb.051994ca

An, H., Roussot, C., Suarez-Lopez, P., Corbesier, L., Vincent, C., Pineiro, M., et al. (2004). CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis. *Development* 131, 3615–3626. doi: 10.1242/dev.01231

Balasubramanian, S., Sureshkumar, S., Agrawal, M., Michael, T. P., Wessinger, C., Maloof, J. N., et al. (2006). The PHYTOCHROME C photoreceptor gene mediates natural variation in flowering and growth responses of Arabidopsis thaliana. *Nat. Genet.* 38, 711–715. doi: 10.1038/ng1818

Bao, S., Hua, C., Huang, G., Cheng, P., Gong, X., Shen, L., et al. (2019). Molecular Basis of Natural Variation in Photoperiodic Flowering Responses. *Dev. Cell* 50, 90–101e103. doi: 10.1016/j.devcel.2019.05.018

Bu, T., Lu, S., Wang, K., Dong, L., Li, S., Xie, Q., et al. (2021). A critical role of the soybean evening complex in the control of photoperiod sensitivity and

adaptation. . *Proc. Natl. Acad. Sci. U S A* 118, e2010241118. doi: 10.1073/pnas.2010241118

Cao, D., Li, Y., Lu, S., Wang, J., Nan, H., Li, X., et al. (2015). GmCOL1a and GmCOL1b Function as Flowering Repressors in Soybean Under Long-Day Conditions. *Plant Cell Physiol.* 56, 2409–2422. doi: 10.1093/pcp/pcv152

Cheng, Z., Zhang, X., Huang, P., Huang, G., Zhu, J., Chen, F., et al. (2020). Nup96 and HOS1 Are Mutually Stabilized and Gate CONSTANS Protein Level, Conferring Long-Day Photoperiodic Flowering Regulation in Arabidopsis. *Plant Cell* 32, 374–391. doi: 10.1105/tpc.19.00661

Choi, I. Y., Hyten, D. L., Matukumalli, L. K., Song, Q., Chaky, J. M., Quigley, C. V., et al. (2007). A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics* 176, 685–696. doi: 10.1534/genetics.107.070821

Dong, L., Fang, C., Cheng, Q., Su, T., Kou, K., Kong, L., et al. (2021b). Genetic basis and adaptation trajectory of soybean from its temperate origin to tropics. *Nat. Commun.* 12:5445. doi: 10.1038/s41467-021-25800-3

Dong, L., Cheng, Q., Fang, C., Kong, L., Yang, H., Hou, Z., et al. (2021a). Parallel selection of distinct Tof5 alleles drove the adaptation of cultivated and wild soybean to high latitude. *Mol. Plant* 15, 308–321. doi: 10.1016/j.molp.2021.10.004

Fan, C., Hu, R., Zhang, X., Wang, X., Zhang, W., Zhang, Q., et al. (2014). Conserved CO-FT regulons contribute to the photoperiod flowering control in soybean. *BMC Plant Biol.* 14:9. doi: 10.1186/1471-2229-14-9

Fang, C., Liu, J., Zhang, T., Su, T., Li, S., Cheng, Q., et al. (2021). A recent retrotransposon insertion of J caused E6 locus facilitating soybean adaptation into low latitude. *J. Integr. Plant Biol.* 63, 995–1003. doi: 10.1111/jipb.13034

Fang, X., Han, Y., Liu, M., Jiang, J., Li, X., Lian, Q., et al. (2021). Modulation of evening complex activity enables north-to-south adaptation of soybean. *Sci. China Life Sci.* 64, 179–195. doi: 10.1007/s11427-020-1832-2

Fehr, W. R., and Caviness, C. E. (1977). *Stages of Soybean Development. Special Report No.80. Cooperative Extension Service, Agriculture and Home Economic Experiment Station Ames*. Iowa: Iowa State University.

Fornara, F., Panigrahi, K. C., Gissot, L., Sauerbrunn, N., Ruhl, M., Jarillo, J. A., et al. (2009). Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev. Cell* 17, 75–86. doi: 10.1016/j.devcel.2009.06.015

Gangappa, S. N., and Botto, J. F. (2014). The BBX family of plant transcription factors. *Trends Plant Sci.* 19, 460–470. doi: 10.1016/j.tplants.2014.01.010

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

Irwin, J. A., Soumpourou, E., Lister, C., Ligthart, J. D., Kennedy, S., and Dean, C. (2016). Nucleotide polymorphism affecting FLC expression underpins heading date variation in horticultural brassicas. *Plant J.* 87, 597–605. doi: 10.1111/tpj.13221

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, H., Jiang, B., Wu, C., Lu, W., Hou, W., Sun, S., et al. (2014). Maturity group classification and maturity locus genotyping of early-maturing soybean varieties from high-latitude cold regions. *PLoS One* 9:e94139. doi: 10.1371/journal.pone.0094139

Jiang, B., Zhang, S., Song, W., Khan, M. A. A., Sun, S., Zhang, C., et al. (2019). Natural variations of FT family genes in soybean varieties covering a wide range of maturity groups. *BMC Genom.* 20:230. doi: 10.1186/s12864-019-5577-5

Kong, F. J., Nan, H. Y., Cao, D., Li, Y., Wu, F. F., Wang, J. L., et al. (2014). A New Dominant Gene E9 Conditions Early Flowering and Maturity in Soybean. *Crop Sci.* 54, 2529–2535.

Langewisch, T., Zhang, H., Vincent, R., Joshi, T., Xu, D., Bilyeu, K., et al. (2014). Major soybean maturity gene haplotypes revealed by SNPViz analysis of 72 sequenced soybean genomes. *PLoS One* 9:e94150. doi: 10.1371/journal.pone.0094150

Laubinger, S., Marchal, V., Le Gourrierec, J., Wenkel, S., Adrian, J., Jang, S., et al. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133, 3213–3222. doi: 10.1242/dev.02481

Lazaro, A., Mouriz, A., Pineiro, M., and Jarillo, J. A. (2015). Red Light-Mediated Degradation of CONSTANS by the E3 Ubiquitin Ligase HOS1 Regulates Photoperiodic Flowering in Arabidopsis. *Plant Cell* 27, 2437–2454. doi: 10.1105/tpc.15.00529

Lazaro, A., Valverde, F., Pineiro, M., and Jarillo, J. A. (2012). The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. *Plant Cell* 24, 982–999. doi: 10.1105/tpc.110.081885

Lestari, P., Lee, G., Ham, T. H., Reflinur Woo, M. O., and Piao, R. (2011). Single nucleotide polymorphisms and haplotype diversity in rice sucrose synthase 3. *J. Hered.* 102, 735–746. doi: 10.1093/jhered/esr094

Li, C., Li, Y. H., Li, Y., Lu, H., Hong, H., Tian, Y., et al. (2020). A Domestication-Associated Gene GmPRR3b Regulates the Circadian Clock and Flowering Time in Soybean. *Mol. Plant* 13, 745–759. doi: 10.1016/j.molp.2020.01.014

Li, J., Wang, X., Song, W., Huang, X., Zhou, J., Zeng, H., et al. (2017). Genetic variation of maturity groups and four E genes in the Chinese soybean mini core collection. *PLoS One* 12:e0172106. doi: 10.1371/journal.pone.0172106

Li, Y. H., Zhang, C., Gao, Z. S., Smulders, M. J. M., Ma, Z. L., Liu, Z. X., et al. (2009). Development of SNP markers and haplotype analysis of the candidate gene for rhg1, which confers resistance to soybean cyst nematode in soybean. *Mol. Breeding* 24, 63–76. doi: 10.3389/fpls.2019.00401

Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., et al. (2013). Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3* 3, 1287–1299. doi: 10.1534/g3.113.006148

Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K., and Abe, J. (2008). Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome A gene. *Genetics* 180, 995–1007. doi: 10.1534/genetics.108.092742

Liu, L. J., Zhang, Y. C., Li, Q. H., Sang, Y., Mao, J., Lian, H. L., et al. (2008). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell* 20, 292–306. doi: 10.1105/tpc.107.057281

Liu, X., Wu, J. A., Ren, H., Qi, Y., Li, C., Cao, J., et al. (2017). Genetic variation of world soybean maturity date and geographic distribution of maturity groups. *Breed. Sci.* 67, 221–232. doi: 10.1270/jsbbs.16167

Lu, S., Dong, L., Fang, C., Liu, S., Kong, L., Cheng, Q., et al. (2020). Stepwise selection on homeologous PRR genes controlling flowering and maturity during soybean domestication. *Nat. Genet.* 52, 428–436. doi: 10.1038/s41588-020-0604-7

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49, 773–779. doi: 10.1038/ng.3819

Patil, G., Do, T., Vuong, T. D., Valliyodan, B., Lee, J. D., Chaudhary, J., et al. (2016). Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. *Sci. Rep.* 6:19199. doi: 10.1038/srep19199

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Robson, F., Costa, M. M., Hepworth, S. R., Vizir, I., Pineiro, M., Reeves, P. H., et al. (2001). Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. *Plant J.* 28, 619–631. doi: 10.1046/j.1365-313x.2001.01163.x

Rosas, U., Mei, Y., Xie, Q., Banta, J. A., Zhou, R. W., Seufferheld, G., et al. (2014). Variation in Arabidopsis flowering time associated with cis-regulatory variation in CONSTANS. *Nat. Commun.* 5:3651. doi: 10.1038/ncomms4651

Samanfar, B., Molnar, S. J., Charette, M., Schoenrock, A., Dehne, F., Golshani, A., et al. (2017). Mapping and identification of a potential candidate gene for a novel maturity locus. *E10, in soybean Theor. Appl. Genet.* 130, 377–390. doi: 10.1007/s00122-016-2819-7

Sawa, M., Nusinow, D. A., Kay, S. A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318, 261–265. doi: 10.1126/science.1146994

Shao, G., Tang, S., Chen, M., Wei, X., He, J., Luo, J., et al. (2013). Haplotype variation at Badh2, the gene determining fragrance in rice. *Genomics* 101, 157–162. doi: 10.1016/j.ygeno.2012.11.010

Strayer, C., Oyama, T., Schultz, T. F., Raman, R., Somers, D. E., Mas, P., et al. (2000). Cloning of the Arabidopsis clock gene TOC1, an autoregulatory response regulator homolog. *Science* 289, 768–771. doi: 10.1126/science.289.5480.768

Suarez-Lopez, P., Wheatley, K., Robson, F., Onouchi, H., Valverde, F., Coupland, G., et al. (2001). CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature* 410, 1116–1120. doi: 10.1038/35074138

Sun, J., Wang, M., Zhao, C., Liu, T., Liu, Z., Fan, Y., et al. (2021). GmFULc Is Induced by Short Days in Soybean and May Accelerate Flowering in Transgenic Arabidopsis thaliana. *Int. J. Mol. Sci.* 22:10333. doi: 10.3390/ijms221910333

Takada, S., and Goto, K. (2003). Terminal flower2, an Arabidopsis homolog of heterochromatin protein1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. *Plant Cell* 15, 2856–2865. doi: 10.1105/tpc.016345

Thakare, D., Kumudini, S., and Dinkins, R. D. (2010). Expression of flowering-time genes in soybean E1 near-isogenic lines under short and long day conditions. *Planta* 231, 951–963. doi: 10.1007/s00425-010-1100-6

Valverde, F., Mouradov, A., Soppe, W., Ravenscroft, D., Samach, A., Coupland, G., et al. (2004). Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science* 303, 1003–1006. doi: 10.1126/science.1091761

Van Inghelandt, D., Melchinger, A. E., Martinant, J. P., and Stich, B. (2012). Genome-wide association mapping of flowering time and northern corn leaf blight (Setosphaeria turcica) resistance in a vast commercial maize germplasm set. *BMC Plant Biol.* 12:56. doi: 10.1186/1471-2229-12-56

Wang, F. F., Nan, H. Y., Chen, L. Y., Fang, C., Zhang, H. Y., Su, T., et al. (2019). A new dominant locus, E11, controls early flowering time and maturity in soybean. *Mol. Breeding* 39:70.

Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., et al. (2020). Natural variation and CRISPR/Cas9-mediated mutation in GmPRR37 affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* 18, 1869–1881. doi: 10.1111/pbi.13346

Wang, X. B., Liu, Z. X., Yang, C. Y., Xu, R., Lu, W. G., Zhang, L. F., et al. (2016). Stability of growth periods traits for soybean cultivars across multiple locations. *J. Integr. Agr.* 15, 963–972. doi: 10.1016/s2095-3119(15)61152-2

Watanabe, S., Hideshima, R., Xia, Z., Tsubokura, Y., Sato, S., Nakamoto, Y., et al. (2009). Map-based cloning of the gene associated with the soybean maturity locus E3. *Genetics* 182, 1251–1262. doi: 10.1534/genetics.108.098772

Watanabe, S., Xia, Z., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., et al. (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. *Genetics* 188, 395–407. doi: 10.1534/genetics.110.125062

Weber, A. L., Zhao, Q., McMullen, M. D., and Doebley, J. F. (2009). Using association mapping in teosinte to investigate the function of maize selection-candidate genes. *PLoS One* 4:e8227. doi: 10.1371/journal.pone.0008227

Wu, C. C., Wei, F. J., Chiou, W. Y., Tsai, Y. C., Wu, H. P., Gotarkar, D., et al. (2020). Studies of rice Hd1 haplotypes worldwide reveal adaptation of flowering time to different environments. *PLoS One* 15:e0239028. doi: 10.1371/journal.pone.0239028

Wu, F., Kang, X., Wang, M., Haider, W., Price, W. B., Hajek, B., et al. (2019). Transcriptome-Enabled Network Inference Revealed the GmCOL1 Feed-Forward Loop and Its Roles in Photoperiodic Flowering of Soybean. *Front. Plant Sci.* 10:1221. doi: 10.3389/fpls.2019.01221

Wu, F., Price, B. W., Haider, W., Seufferheld, G., Nelson, R., and Hanzawa, Y. (2014). Functional and evolutionary characterization of the CONSTANS gene family in short-day photoperiodic flowering in soybean. *PLoS One* 9:e85754. doi: 10.1371/journal.pone.0085754

Xia, Z., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., et al. (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. U S A* 109, E2155–E2164. doi: 10.1073/pnas.1117982109

Yonemaru, J., Ebana, K., and Yano, M. (2014). HapRice, an SNP haplotype database and a web tool for rice. *Plant Cell Physiol.* 55, e9. doi: 10.1093/pcp/pct188

Yonemaru, J., Yamamoto, T., Ebana, K., Yamamoto, E., Nagasaki, H., Shibaya, T., et al. (2012). Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. *PLoS One* 7:e32982. doi: 10.1371/journal.pone.0032982

Yue, Y., Liu, N., Jiang, B., Li, M., Wang, H., Jiang, Z., et al. (2017). A Single Nucleotide Deletion in J Encoding GmELF3 Confers Long Juvenility and Is Associated with Adaptation of Tropic Soybean. *Mol. Plant* 10, 656–658. doi: 10.1016/j.molp.2016.12.004

Yue, Y. L., Sun, S., Li, J. W., Yu, H. D., Wu, H. X., Sun, B. Q., et al. (2021). GmFULa improves soybean yield by enhancing carbon assimilation without altering flowering time or maturity. *Plant Cell Rep.* 40, 1875–1888. doi: 10.1007/s00299-021-02752-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

ORIGINAL RESEARCH
published: 22 March 2022
doi: 10.3389/fpls.2022.827701

Hybrids Provide More Options for Fine-Tuning Flowering Time Responses of Winter Barley

Miriam Fernández-Calleja[1], Francisco J. Ciudad[2], Ana M. Casas[1] and Ernesto Igartua[1]*

[1] Department of Genetics and Plant Production, Aula Dei Experimental Station - Spanish National Research Council (EEAD-CSIC), Zaragoza, Spain, [2] Agricultural Technology Institute of Castilla and León (ITACYL), Valladolid, Spain

Crop adaptation requires matching resource availability to plant development. Tight coordination of the plant cycle with prevailing environmental conditions is crucial to maximizing yield. It is expected that winters in temperate areas will become warmer, so the vernalization requirements of current cultivars can be desynchronized with the environment's vernalizing potential. Therefore, current phenological ideotypes may not be optimum for future climatic conditions. Major genes conferring vernalization sensitivity and phenological responses in barley (*Hordeum vul*gare L.) are known, but some allelic combinations remain insufficiently evaluated. Furthermore, there is a lack of knowledge about flowering time in a hybrid context. To honor the promise of increased yield potentials, hybrid barley phenology must be studied, and the knowledge deployed in new cultivars. A set of three male and two female barley lines, as well as their six F_1 hybrids, were studied in growth chambers, subjected to three vernalization treatments: complete (8 weeks), moderate (4 weeks), and low (2 weeks). Development was recorded up to flowering, and expression of major genes was assayed at key stages. We observed a gradation in responses to vernalization, mostly additive, concentrated in the phase until the initiation of stem elongation, and proportional to the allele constitution and dosage present in *VRN-H1*. These responses were further modulated by the presence of *PPD-H2*. The duration of the late reproductive phase presented more dominance toward earliness and was affected by the rich variety of alleles at *VRN-H3*. Our results provide further opportunities for fine-tuning total and phasal growth duration in hybrid barley, beyond what is currently feasible in inbred cultivars.

Keywords: vernalization sensitivity, adaptation, gene action, hybrid breeding, preanthesis phenological phases

OPEN ACCESS

Edited by:
Yang Zhu,
Zhejiang University, China

Reviewed by:
Zhi-Ming Yu,
Hangzhou Normal University, China
Mingxun Chen,
Northwest A&F University, China

****Correspondence:***
Ana M. Casas
acasas@eead.csic.es

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

Received: *02 December 2021*
Accepted: *07 February 2022*
Published: *22 March 2022*

Citation:
Fernández-Calleja M, Ciudad FJ, Casas AM and Igartua E (2022) Hybrids Provide More Options for Fine-Tuning Flowering Time Responses of Winter Barley. Front. Plant Sci. 13:827701. doi: 10.3389/fpls.2022.827701

INTRODUCTION

Higher and more stable crop yields are the main targets for cereal breeders. This goal is increasingly challenging in temperate regions, where major crops face growing threats from the impact of climate change, particularly from drought and heat events at critical developmental milestones during the crop cycle (Olesen et al., 2010; Porter et al., 2014; Trnka et al., 2014). Tight coordination of plant cycle to environmental conditions to match resource availability with the most sensitive

growth stages is crucial for crop adaptation (Craufurd and Wheeler, 2009), and has a major effect on yield (Bolaños and Edmeades, 1993; Evans, 1996; González et al., 1999; Cockram et al., 2007; Tondelli et al., 2014; Flohr et al., 2018; Wiegmann et al., 2019). In this context, the current variety formats for cultivation should be re-assessed, as they may no longer be the highest-yielding ones. Further research on crop plasticity is necessary to adapt cereal crops to the range of future climatic conditions (Fatima et al., 2020). Winters in the temperate zone are projected to be warmer, so the vernalization requirement of current winter cultivars may be excessive, i.e., may not be met on time, due to a lower vernalizing potential of the environment (Saadi et al., 2015; Yang et al., 2019). Future ideotypes will have to combine specific vernalization and photoperiod responses fine-tuned to the projected climatic conditions prevalent for each region (Stratonovitch and Semenov, 2015; Gouache et al., 2017; Tao et al., 2017). Allelic variation at the *VRN-H1* gene already induces a gradation of vernalization needs to the barley plants, which have had a large impact on barley adaptation to regional climates (Casao et al., 2011b; Contreras-Moreira et al., 2019). Breeders must aim at deploying appropriate phenology gene combinations to optimize the crop foundation phase (vegetative and early reproductive), and construction phase (late reproductive) growth periods, as well as avoiding abiotic stresses at critical developmental stages, thus optimizing yield potential in target environments (Gouache et al., 2017).

Nowadays, there is growing interest in breeding hybrid cereal varieties, including barley. Hybrids have shown greater yield potential than inbred lines, due to exploitation of heterosis, greater yield stability under fluctuating environmental conditions, and the ease of pyramiding strategic combinations of dominant major genes (Longin et al., 2012; Mühleisen et al., 2013, 2014). Therefore, optimizing phenology in hybrid cultivars is a strategy to improve yields under current and future climate conditions. However, there is a lack of knowledge about flowering time gene action in a hybrid context.

Flowering time in barley is tightly regulated by genetic networks that respond predominately to day-length (photoperiod) and prolonged exposure to cold temperature (vernalization). Barley is a facultative long-day plant, flowering earlier under increasing day-lengths, and characterized by two major growth types, namely, winter and spring. Winter barleys need vernalization for timely flowering (Campoli and von Korff, 2014).

Vernalization genetic control is based on the epistatic system composed of flowering inducer *VRN-H1* (Trevaskis et al., 2003; Yan et al., 2003) and repressor *VRN-H2* (Yan et al., 2004). *VRN-H1* corresponds to gene *HvBM5A*, an orthologue of MADS-box *AP1* from *Arabidopsis* (Trevaskis et al., 2007), whereas *VRN-H2* has no clear correspondence in *Arabidopsis*. Winter barleys carry the functional dominant *VRN-H2* allele, accompanied by a cold-sensitive *VRN-H1* allele. Activation of *VRN-H1* is quantitative, with long cold treatments inducing higher levels of expression (von Zitzewitz et al., 2005; Sasani et al., 2009), which results in an earlier transition to the reproductive phase (Sasani et al., 2009). It presents a large number of alleles, which are defined by the length of the first intron (11 kb in the wild-type *vrn-H1*), and present a gradation of responses to vernalization (Takahashi and Yasuda, 1971; Szűcs et al., 2007), roughly proportional to the first intron length (Szűcs et al., 2007; Hemming et al., 2009; Casao et al., 2011a; Oliver et al., 2013; Guerra et al., 2021). Regarding the gene action of the *VRN-H1* allelic series, the accepted model states that the winter allele is recessive, while the rest are dominant (Takahashi and Yasuda, 1971; Haas et al., 2020). *VRN-H3* (*HvFT1*) is the key flowering inducer that integrates the photoperiod and vernalization pathways (Yan et al., 2006; Faure et al., 2007; Kikuchi et al., 2009), whose expression is induced under long-day conditions and promotes flowering (Turner et al., 2005; Hemming et al., 2008), and is an orthologue of *FT1* in *Arabidopsis* (Yan et al., 2006). Ample allelic variation at *VRN-H3* has been described, arising from sequence polymorphisms in the promoter and first intron (Yan et al., 2006; Hemming et al., 2008; Casas et al., 2011, 2021), and copy number variation (Nitcher et al., 2013; Loscos et al., 2014), as summarized in Fernández-Calleja et al. (2021), but there is no information on its gene action. According to the currently accepted model, during autumn, when temperate cereals germinate, *VRN-H2* represses *VRN-H3* expression. During winter, vernalization induces *VRN-H1* expression, resulting in *VRN-H2* repression in leaves and, consequently, activation of *VRN-H3* transcription in spring, which promotes the transition from the vegetative to the reproductive stage (Trevaskis et al., 2006; Distelfeld et al., 2009). At the whole plant level, this transition is visible as the appearance of the first node at the main stem, and the beginning of stem elongation (jointing stage). Besides the *VRN* genes, *HvODDSOC2* also plays a repressor role in the vernalization pathway (Greenup et al., 2010). This gene is the monocot orthologue of *Arabidopsis thaliana FLOWERING LOCUS C* (*FLC*, Ruelens et al., 2013). It is downregulated by prolonged cold exposure, was identified as a binding target of the VRN1 protein in barley, together with *VRN-H2* and *VRN-H3* genes (Deng et al., 2015), and plays a repressor role in absence of full vernalization (Monteagudo et al., 2019b). Genes *PPD-H1* and *PPD-H2* control photoperiod sensitivity. *PPD-H1* (*HvPRR37*, orthologue of *PRR7* in *Arabidopsis*) is the major determinant of long photoperiod response in barley (Turner et al., 2005). Its activity causes an increased expression of *VRN-H3* after vernalization fulfillment, promoting flowering under long-day conditions (Turner et al., 2005; Campoli et al., 2012; Mulki and von Korff, 2016). *PPD-H2* (*HvFT3*), which belongs to the FT gene family, induces early reproductive development in short-day conditions, or even long-day conditions when vernalization requirements have not been fully satisfied (Laurie et al., 1995; Faure et al., 2007; Casao et al., 2011a,b; Mulki et al., 2018). Phylogeographic and genetic studies suggest an adaptive role for this gene in winter barleys (Kikuchi et al., 2009; Casao et al., 2011b).

Barley breeding for the near future requires understanding the genetic mechanisms of adaptation, including vernalization responses, in a hybrid context. This work is intended to explore the inheritance and effect of several major flowering genes in heterozygosis, and their dynamics in relation to insufficient vernalization. For this purpose, an experiment with different vernalization treatments was designed aiming to evaluate the phenology and gene expression of key genes in the plant cycle

duration, in a set of hybrid barleys and their parents. Here we show that hybrid combinations extend the available catalog of genetic responses to vernalization, opening new possibilities for optimizing phenology to specific areas using hybrids.

MATERIALS AND METHODS

Plant Material

In this study, eleven barley (*Hordeum vulgare* L.) genotypes were used: two female parents (Female A and Female B), three pollinators (Male 1, Male 2, and Male 3), and six hybrids derived from their crosses (Hybrid A1, Hybrid A2, Hybrid A3, Hybrid B1, Hybrid B2, and Hybrid B3). The female parents are cytoplasmic male sterile (CMS) inbred lines used in the development of 6-row winter barley hybrids for Europe by Syngenta® (Basel, Switzerland). The pollinators are advanced inbred lines developed in the framework of the Spanish Barley Breeding Program (Gracia et al., 2012), well adapted to the Mediterranean conditions, and without fertility restorer genes. The resultant offspring are male-sterile hybrids (female hybrids, F_1F), an intermediate step in the production of a three-way hybrid, after further crossing with a fertility restorer genotype. The genotypes studied present different *VRN-H1* alleles, which are defined by the length of the first intron of the gene, and have different vernalization requirements, ranking from low to high cold needs. These alleles are *VRN-H1-4*, which presents a vernalization requirement of around 2 weeks, *VRN-H1-6* requires approximately 30 days, and *vrn-H1* requires not less than 7 weeks (Casao et al., 2011a). These alleles represent the main allelic diversity that is spread across Western European six-row winter barley. The strict winter allele (*vrn-H1*) prevails in North-western European barleys, whereas alleles *VRN-H1-4* and *VRN-H1-6* correspond to the two largest germplasm groups found in Spanish barley landraces (Casao et al., 2011a). The geographical distribution of these two alleles coincides with the harshness of winters in the Iberian region (Yahiaoui et al., 2008; Casao et al., 2011a; Contreras-Moreira et al., 2019). *VRN-H1-4* predominates in Southern and coastal Spanish landraces, while *VRN-H1-6* is frequently found in the continental inlands of Spain. Besides, the genotypes studied also present different alleles at other major genes involved in the control of vernalization responses and day-length sensitivity (**Table 1**).

TABLE 1 | Genotypes for the genes associated with responses to vernalization and photoperiod in the parent lines under study.

	Vernalization, photoperiod, and *earliness per se* genes				
Parent lines	***VRN-H1*[a]**	***VRN-H2*[b]**	***VRN-H3*[c]**	***PPD-H1*[d]**	***PPD-H2*[e]**
Female A	*vrn-H1*	*VRN-H2*	*vrn-H3d(1)*	*PPD-H1*	*ppd-H2*
Female B	*VRN-H1-6*	*VRN-H2*	*vrn-H3a(1)*	*PPD-H1*	*ppd-H2*
Male 1	*VRN-H1-4*	*VRN-H2*	*vrn-H3d(1)*	*PPD-H1*	*ppd-H2*
Male 2	*VRN-H1-6*	*VRN-H2*	*vrn-H3c(1)*	*PPD-H1*	*ppd-H2*
Male 3	*vrn-H1*	*VRN-H2*	*vrn-H3a(1)*	*PPD-H1*	*PPD-H2*

[a] *Alleles based on the size of intron 1, following Hemming et al. (2009).*
[b] *Presence (dominant)/absence (recessive) of HvZCCT, following Karsai et al. (2005).*
[c] *Alleles based on two indels in the first 550 bp upstream of the start codon, two SNPs in intron 1, and CNV, coded as in Fernández-Calleja et al. (2021). vrn-H3a(1) = promoter deletion-insertion, intron 1 AG, CNV 1 copy, vrn-H3c(1) = deletion-insertion, TC, 1 copy, vrn-H3d(1) = insertion-deletion, TC, 1 copy.*
[d] *Alleles based on SNP22 of Turner et al. (2005), dominant = G, recessive = T.*
[e] *Presence (dominant)/absence (recessive), as in Faure et al. (2007).*

Plant Growth Conditions, Phenotyping, and Sampling

A study under controlled conditions was designed to assess differences in development and gene expression in key genes controlling the duration of the plant cycle, after three vernalization treatments: complete (8 weeks of cold period, V8), moderate (4 weeks, V4), and low (2 weeks, V2).

Genotypes were exposed to treatments of 14 (low), 28 (intermediate), and 58 days (full vernalization) at 6 ± 2°C, under a short-day regime (8 h light/16 h dark). After vernalization, seedlings were transferred to a growth chamber with conditions set to long photoperiod (16 h light/8 h night), 220 μmol m^{-2} s^{-1} light intensity, and 20°C day/16°C night temperatures. The duration of the vernalization treatments was set according to previous experiments. Fourteen days are enough for genotypes carrying the *VRN-H1-4* allele, 58 days are sufficient for cultivars with strict winter growth habits, and 28 days is an intermediate condition.

The study was carried out in two stages, growing plants independently for gene expression and phenotyping, and using the same growth chamber. For phenotyping, plants were grown in trays of 12 cells (650 cc per cell) from sowing until flowering. The number of days to the appearance of the first node at the base of the main stem, or stage 31 on the Zadoks scale (Z31), and the number of days to awn tipping or Z49 (Zadoks et al., 1974) were recorded. Four plants were assessed for each genotype and treatment. After discarding dead plants and outliers, data from the best three plants per treatment were kept. The phenotyping experiment lasted 130 days.

The gene expression experiment was carried out in two batches due to growth chamber capacity, split according to the females, due to space limitations. One subset comprised the hybrids derived from Female A and respective parents (Batch A), and the other subset included the hybrids coming from Female B and respective parents (Batch B). Therefore, the three male parents were present in the two batches. Plants were grown in trays of 35 cells (200 cc per cell), from sowing until sampling. In the second stage (phenotyping), undisturbed plants from all eleven genotypes were simultaneously assessed for developmental traits.

The last expanded leaf of four independent plants was sampled 17 and 35 days after the entry in the growth chamber (i.e., after the end of the respective vernalization treatment), 14 h into the light period (2 h before the end of the day). Samples were frozen in liquid nitrogen, homogenized (Mixer Mill model MM400, Retsch, Haan, Germany) and conserved at −80°C until RNA isolation. In principle, three plants were analyzed. If they were

clearly dissimilar, the fourth plant was also analyzed, and the best three were kept for further analysis.

Gene Expression Analysis

Using the Total RNA Mini Kit for Plants (IBI Scientific, Dubuque, IA, United States) RNA extraction was carried out following manufacturer instructions. Total RNA (1 μg) was employed for cDNA synthesis using SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA, United States) and oligo (dT) 20 primer (Invitrogen, Carlsbad, CA, United States). RT PCR quantification (ABI 7500, Applied Biosystems, Waltham, MA, United States) was performed for samples from each time point and vernalization treatment. Four plants (biological replicates) per sampling time, treatment, and genotype were sampled. Three biological replicates and two technical replicates were tested per sample and pair of primers (*VRN-H1*, *VRN-H2*, *VRN-H3*, *PPD-H1*, *HvODDSOC2*, and *PPD-H2*). When outliers were detected, the fourth plant was also analyzed, and the three replicates which showed the best agreement were kept. Primer sequences and conditions are specified in **Supplementary Table 1**. Gene expression levels were normalized to *Actin* expression, considering primer efficiencies. The average of the two technical replications of ΔCt (Ct *Actin*–Ct target gene) was used as the experimental unit for statistical analyses, to protect against small pipetting errors.

Statistical Analysis

Statistical analyses were carried out using R software (R Core Team, 2013). Differences in Z31, Z49, and the lag between the latter stages (Lag Z31–Z49) between genotypes and treatments were evaluated using the ANOVA procedure in R. The ANOVA model included genotype, vernalization treatment, and genotype by treatment interaction, all taken as fixed factors. The three plants sampled per genotype were considered biological replicates. The genotype factor was broken down into the most informative contrasts: hybrids *vs.* parents, females *vs.* males, and finally, hybrids from Female A *vs.* hybrids from Female B. Multiple comparisons were obtained by Fisher's protected least significant differences (LSD) with the R package "emmeans" (Lenth et al., 2018).

Using the same statistical procedure, differences in vernalization sensitivity for Z31, Z49 and lag Z31–Z49 between genotypes and treatments (V8–V4, V4–V2, and V8–V2) were tested. Vernalization sensitivity comparing V8 and V4 treatments was calculated by subtracting the value for an individual observation in the 8-week vernalization treatment from the mean value of that genotype in the 4-week vernalization treatment, for each suitable variable. The same procedure was followed for each variable considering the V4–V2 or V8–V2 treatments.

For gene expression results, Batch A and Batch B were analyzed separately. The ANOVA model included genotype, treatment, sampling time, and factorial interactions. The analyses of variance were performed considering all factors (genotype, sampling time, and treatment) as fixed. The three biological replications were considered replicates. The contrasts defined were hybrids *vs.* parents, hybrids *vs.* females, hybrids *vs.* males, and females *vs.* males. Means were compared using the LSD test ($P < 0.05$).

A correlation network analysis was carried out with the R package "qgraph" (Epskamp et al., 2012). We performed a multiple factorial analysis (Pagès, 2002) using R packages "FactoMineR" (Lê et al., 2008) and "factoextra" (Kassambara and Mundt, 2017). This method summarizes and displays a complex data table in which individuals are described by several sets of variables (quantitative and /or qualitative) structured into groups. It requires balancing the influences of each set of variables. Therefore, the variables are weighted during the analysis. Variables in the same group are normalized using the same weighting value, which can vary from one group to another. In our case, we summarized the observations described by a set of variables structured into four groups (*Treatment*, *Genotype*, *Development,* and *Gene expression*). *Genotype* and *treatment* are groups based on categorical variables specifying the genotype identity of each individual and the vernalization treatment to which they were subjected. *Development* and *gene expression* quantitative variables were considered as active groups and their contribution was considered to define the distance between individuals. Each variable within a group was equally weighted, so the influence of each set of variables in the analysis was balanced.

RESULTS

Insufficient vernalization markedly extended the growth cycle of plants. At the complete vernalization treatment (V8), the duration of the two phases considered (time until Z31 and lag Z31–Z49) was rather similar. With increasingly insufficient vernalization, time to awn tipping (Z49) raised progressively. Most of this lengthening occurred in the period until the first node appearance (Z31), although additional delays were observed in the late reproductive phase (lag Z31–Z49), particularly for the female genotypes (**Figures 1**, **2**).

Vernalization Response of Parent and Hybrid Genotypes

Differences in the length of developmental phases between genotypes were also detected (**Supplementary Tables 2, 6**). Male parents (the Mediterranean adapted) were earlier than female parents. In general, hybrids showed an intermediate phenotype between both parents for days to reach the jointing stage (Z31), and days to awn tipping in all vernalization treatments. However, the length of the late reproductive phase of the hybrids was closer to that of the male parents (**Figures 1**, **2** and **Supplementary Table 3**).

Parents and hybrids presented different vernalization responses (**Figure 2**). This differential behavior was mostly explained by changes in the duration of the phase until jointing, which was associated with the *VRN-H1* allele present (negative correlation indicated in **Supplementary Figure 1**), and their dosage in the case of the hybrids. Increasing vernalization treatments minimized these differences between genotypes, especially in those triads (female, hybrid, male) where different vernalization alleles were crossed (**Figure 2**).

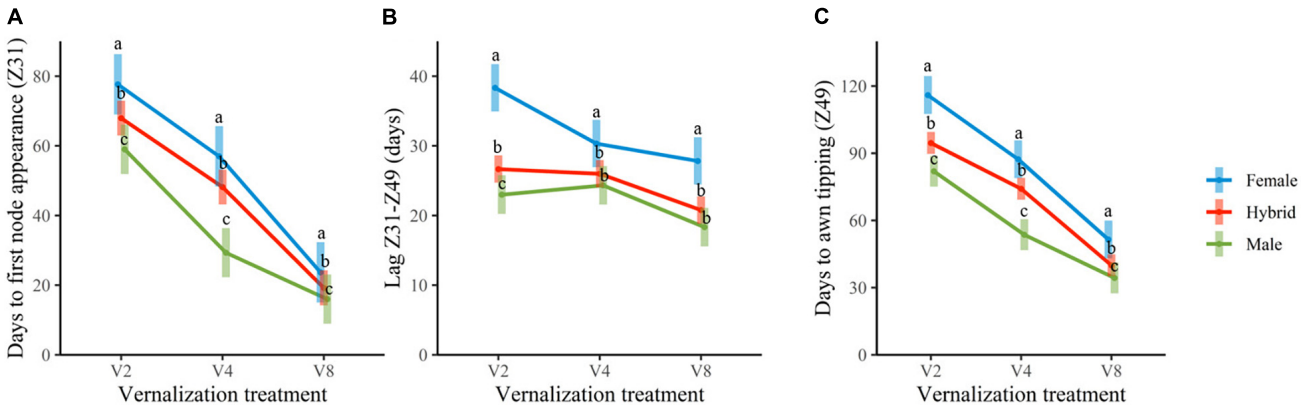


FIGURE 1 | From left to right, the average duration of time until Z31, **(A)**, lag Z31–Z49 **(B)**, and total time until Z49 **(C)**, for the three groups of genotypes, male parents, female parents, and hybrids. Error bars are 95% CIs. For each developmental variable and treatment, group means with a different letter are significantly different at $P < 0.05$ according to the contrasts performed for the overall ANOVA.

With complete vernalization (V8), all genotypes reached the Z31 stage in a similar range, although some differences were still evident. Male 1 and Male 2, carrying low (*VRN-H1-4*) and medium (*VRN-H1-6*) vernalization requirement alleles, were the earliest until Z31. Winter parents (*vrn-H1*), Male 3 and Female A, and Female B (*VRN-H1-6*), showed intrinsic lateness, reaching the Z31 stage 10 days later than the other male parents did. All six hybrids reached first node appearance at approximately the same time, independently of their *VRN-H1* allele, and significantly earlier than their female parents (**Figure 2**).

When vernalization was moderate (V4), the differences in days to first node appearance between genotypes were accentuated (**Figure 2**). Male 1 (*VRN-H1-4*) maintained a short Z31 phase, only 8 days longer compared to V8. Male 2 (*VRN-H1-6*) and Male 3 (*vrn-H1*) were more affected by the reduction of the cold treatment, although they only delayed their Z31 date around 15 days compared to the V8 treatment. The female parents, in contrast, experienced a remarkable delay in days to first node appearance, particularly large for the winter Female A (40 days, **Supplementary Figure 2**). Interestingly, both Male 3 and Female A carry the strict winter *vrn-H1* allele, but they differed almost 30 days at Z31 for the V4 treatment (**Figure 2**). This finding indicates that early alleles of Male 3 at genes other than *VRN-H1* are having a shortening effect on the Z31 phase. This gene could be *PPD-H2*, as will be discussed later. For hybrids, in general, we observed intermediate phenotypes between the behaviors of their parents. Hybrids A1 and B1, carrying one copy of the *VRN-H1-4* allele, were the least affected by the reduction in the cold treatment, reaching Z31 earlier than any other hybrid. Hybrids from Male 2 and Male 3 showed a longer delay in Z31 with reduced vernalization (**Figure 2**).

When vernalization was low (V2), the Z31 phase was prolonged by 15 days or more for most genotypes (compared to V4). The exception was the genotypes carrying the *VRN-H1-4* allele, which reached Z31 considerably earlier than the rest of the genotypes (**Figure 2**) and had a low vernalization sensitivity (**Supplementary Figure 2**). Winter genotypes suffered the largest changes in time until Z31 when comparing V4 and V2, particularly Male 3 and its crosses. The differences between the females were even more pronounced in this treatment (V2), Female A reached Z31 30 days later than Female B (**Figure 2**). This agrees with the alleles that these genotypes carry at *VRN-H1*. Female A carries the winter allele (*vrn-H1*), characterized by a higher sensitivity to vernalization than the *VRN-H1-6* allele of female B. This dissimilarity in Z31 dates between females translated into differences in Z31 duration also between A and B hybrids (**Figure 2**). For instance, Hybrid A2 (*vrn-H1*/*VRN-H1-6*) delayed the jointing stage 16 days more than Hybrid B2 (*VRN-H1-6*/*VRN-H1-6*) when reducing the vernalization treatment from 4 to 2 weeks (**Supplementary Figure 2**), indicating a dosage effect of winter *VRN-H1* alleles. Particularly interesting in this treatment is the change in ranking observed in the slower developing parents. Male 3, which bears the active allele at *PPD-H2*, reached Z31 earlier than both female parents under moderate (V4) to complete (V8) vernalization. By contrast, at V2, Male 3 (*vrn-H1*) delayed significantly its early development, reaching the Z31 stage later than Female B (*VRN-H1-6*) and almost at the same time as Female A (*vrn-H1*) (**Figure 2**). The striking reduction in time to jointing of Male 3 and Hybrid B3 with moderate vernalization, but not in the low vernalization treatment, agrees well with the hypothesis that *PPD-H2* needs some cold to come into play (Monteagudo et al., 2019b).

In summary, the Z31 phase was clearly the most sensitive period to the cold treatment. Vernalization sensitivity differed between genotypes (**Supplementary Tables 2, 7**), and seemed related to the *VRN-H1* allele present and proportional to the *VRN-H1* allele dosage. Genotypes carrying the low vernalization requirement allele *VRN-H1-4* reached the jointing stage earlier than the rest of the genotypes regardless of the vernalization treatment. In contrast, genotypes carrying winter *vrn-H1* alleles increased steeply the time to reach Z31 stage in the low vernalization treatment. Genotypes with a *VRN-H1-6* allele delayed jointing stage under insufficient vernalization, but not as much as strict winter types. We recoded *VRN-H1* alleles as a

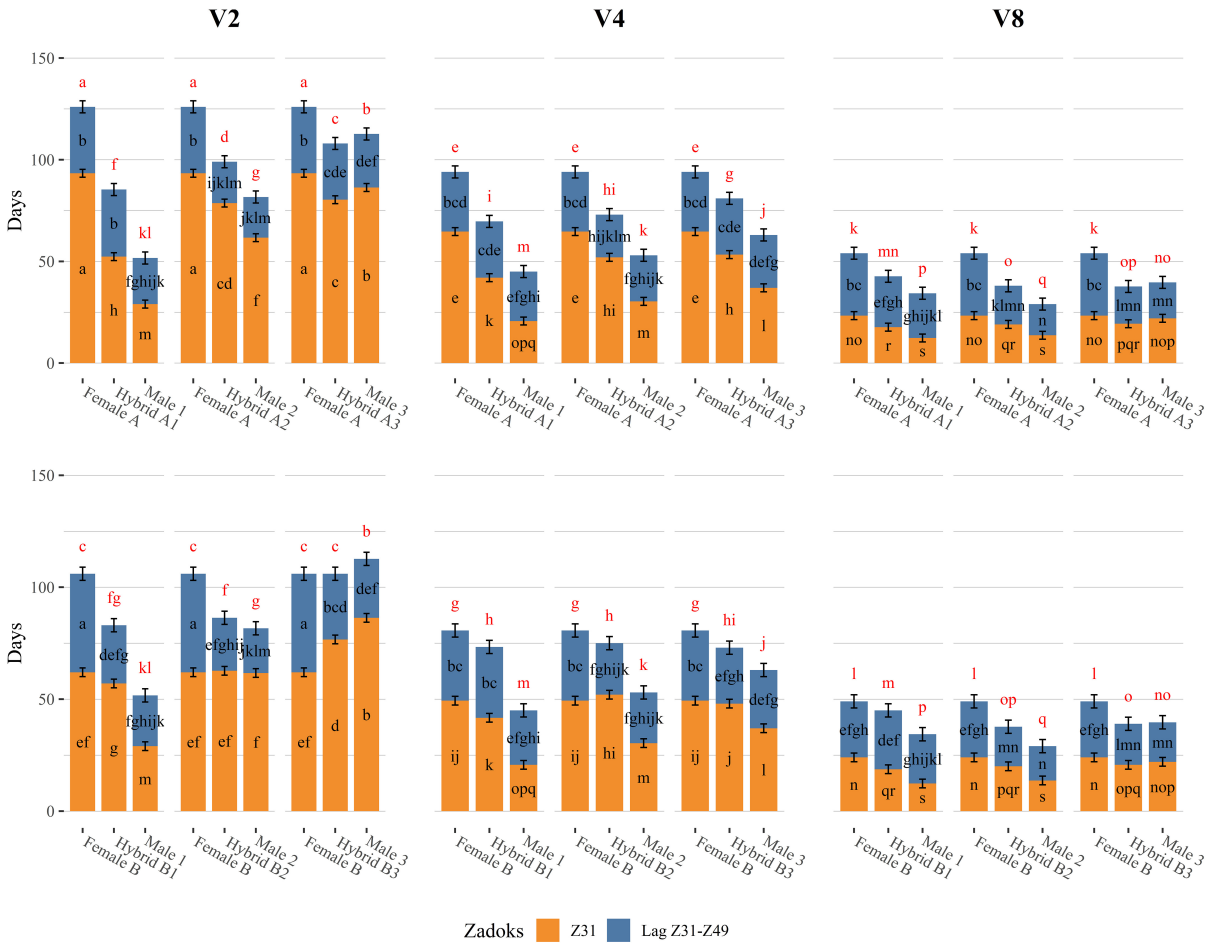


FIGURE 2 | Developmental differences of triads of genotypes (Female, Hybrid, and Male) grown under 16 h light, in response to different vernalization treatments (V2: 2 weeks of vernalization, V4: 4 weeks of vernalization, V8: 8 weeks of vernalization; 4–8°C, 8 h light). Days to first node appearance (Z31) are represented as the height of the orange bars, days to awn tipping (Z49) are represented as the total height of the bars, and the lag period between Z31 and Z49 (LagZ31–Z49) is represented as the segment in blue. Genotypes from Batch A are shown in the upper part of the graph, whereas genotypes from Batch B are shown in the bottom part of the graph. Each of the three columns of facets represents a vernalization treatment, low on the left, moderate in the middle, and complete on the right. Each subplot within each facet contains one triad of genotypes composed of one female parent on the left, the male parent on the right, and the hybrid in the middle. Error bars are 95% CIs for Z31 and LagZ31–Z49. For each developmental variable, bars with a different letter are significantly different at $P < 0.05$. Letters in red represent the comparison of the means for Z49.

categorical variable with integer numbers roughly proportional to their associated vernalization response (1-2-3, for *VRN-H1-4*, *VRNH1-6*, and *vrn-H1*, respectively). We calculated a synthetic vernalization score by adding the values for the two alleles carried by each genotype. A regression of vernalization sensitivity (in this case, the difference V8–V2 for Z31) on the genotypic score produced a very good fit (**Figure 3**), supporting the dosage effect of *VRN-H1*.

The differences observed for the jointing phase were maintained until awn tipping but modulated by the effect of other genes on the late reproductive phase, likely *VRN-H3* and *PPD-H2* (**Supplementary Figure 1**). Male 1 and its crosses showed a constant and rather long lag Z31–Z49 phase across cold treatments. Conversely, Male 2 and its hybrids stood out for a consistent short late reproductive phase across treatments, probably because they carry the fast *vrn-H3c(1)* allele at *VRN-H3*. This additional precocity allowed them to be the fastest genotypes in reaching Z49 in V8, among parents and hybrids, respectively, even earlier than *VRN-H1-4* carriers. Male 3 and its crosses also showed a particularly short lag Z31–Z49 phase, but only at V8. Interestingly, female parents differed in their late reproductive phase duration patterns. Female A showed a constant duration of the late reproductive phase across treatments, whereas Female B showed a lag Z31–Z49 phase increasingly long with decreasing vernalization treatments (**Figure 2**). This observation was not translated into the hybrids, as they more closely resembled their male parents in the duration of the late reproductive phase.

Focusing on inheritance, hybrids derived from Male 1 and Male 2 showed an intermediate phenotype for Z31 and Z49, between the early male parents and the late female parents, in all vernalization treatments. Nevertheless, the lag Z31–Z49 phase of Male 2 hybrids was as short as that of their male parent

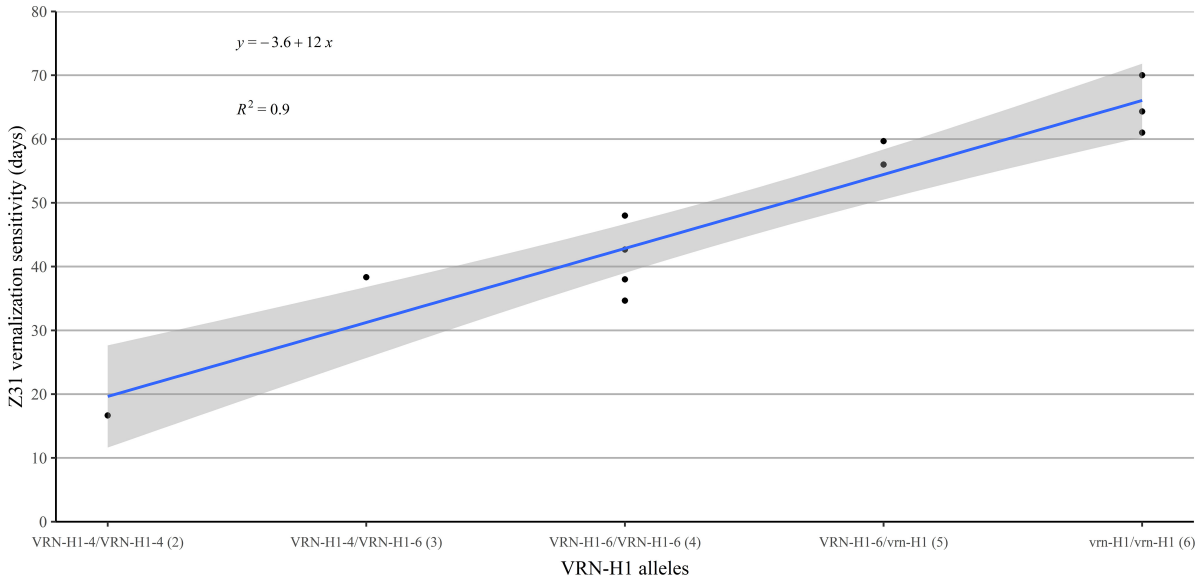


FIGURE 3 | Regression of vernalization sensitivity (duration until Z31 at V2 minus duration until Z31 in V8) on the allelic constitution at *VRN-H1*, coded as 1 (*VRN-H1-4*), 2 (*VRN-H1-6*), or 3 (*vrn-H1*), proportional to the vernalization requirement induced by each allele. The number within brackets indicates the vernalization requirement score of the genotype, according to their *VRN-H1* alleles.

(**Figure 2**), indicating the dominance of the Male 2 early allele controlling this phase. Hybrids A3 and B3, in contrast to the other hybrids, did not show a consistent intermediate phenotype between their parents. In the low and complete vernalization treatments, hybrids A3 and B3 headed Z49 as early as the early parent, due to a dominant short late reproductive phase similar to that of Male 3 (**Figure 2**).

Gene Expression

Differences among genotypes were detected for the expression of all genes tested (**Supplementary Tables 4, 5, 8, 9**).

In all genotypes, *VRN-H1* expression increased gradually with increasing duration of vernalization (**Figure 4** and **Supplementary Figure 3**), although differences between *VRN-H1* alleles were evident. Male 1 and its hybrids, carrying *VRN-H1-4*, were the only genotypes showing upregulated *VRN-H1* expression after just 2 weeks of vernalization. After vernalization for 4 weeks, *VRN-H1* expression also reached high levels in Male 2 and Hybrid B2, both carrying the medium vernalization requirement allele *VRN-H1-6*. The rest of the parents and hybrids required 8 weeks of cold to show high *VRN-H1* transcript levels (**Figures 4A,D**). However, not all differences in *VRN-H1* expression levels were due to allelic differences. For instance, both Female B and Male 2 carry the *VRN-H1-6* allele but present different *VRN-H1* expressions at V4 (**Figures 4A,D**). Hybrid B2 had the same (higher) *VRN-H1* expression as Male 2, indicating higher repression of *VRN-H1* in Female B, which is lost in the hybrid. In general, *VRN-H1* expression levels paralleled plant development patterns across genotypes and treatments.

When vernalization was complete, all parents and hybrids showed a high *VRN-H1* expression, probably caused by saturating vernalization requirements. With incomplete vernalization, however, *VRN-H1* expression in hybrids was, in general, intermediate between their parents (**Figures 4A,D**). Additive gene action caused by a dosage effect of *VRN-H1* was visible at the gene expression level, supporting the hypothesis that the duration of the phase until jointing is related to the effect of this gene. This was apparent when comparing Hybrid A2 (*vrn-H1*/*VRN-H1-6*), which did not peak until V8 (**Figure 4A**), with Hybrid B2 (*VRN-H1-6*/*VRN-H1-6*), in which *VRN-H1* expression peaked already with 4 weeks of vernalization (**Figure 4D**).

All lines carried the active *VRN-H2* allele, but differences in its expression were observed (**Figures 4B,E**). As expected, *VRN-H2* expression decreased with increasing duration of the vernalization treatment, and inversely correlated to the expression of *VRN-H1*. A similar trend was observed for *HvODDSOC2* (**Supplementary Figures 4B,E**). Nevertheless, there were some differences in *VRN-H2* expression among genotypes carrying the same winter *VRN-H1* allele. Indeed, at V4 and V8 treatments, the repression of *VRN-H2* in Male 3 and Hybrid A3 was higher than in the Female A (**Figure 4B**), despite all being winter types. This result agrees with an antagonistic relationship between *VRN-H2* and *PPD-H2* (present only in Male 3 and its hybrids).

With increasing duration of vernalization (**Figures 4C,F**), *VRN-H3* expression increased, and followed closely that of *VRN-H1*. There were no apparent differences in expression between *VRN-H3* alleles. Again, we could observe differences in expression among the two winter parents (Female A and Male 3), with Male 3 showing the highest *VRN-H3* expression in all vernalization treatments.

Across vernalization treatments and genotypes, *PPD-H1* expression was consistently high. All genotypes assessed in

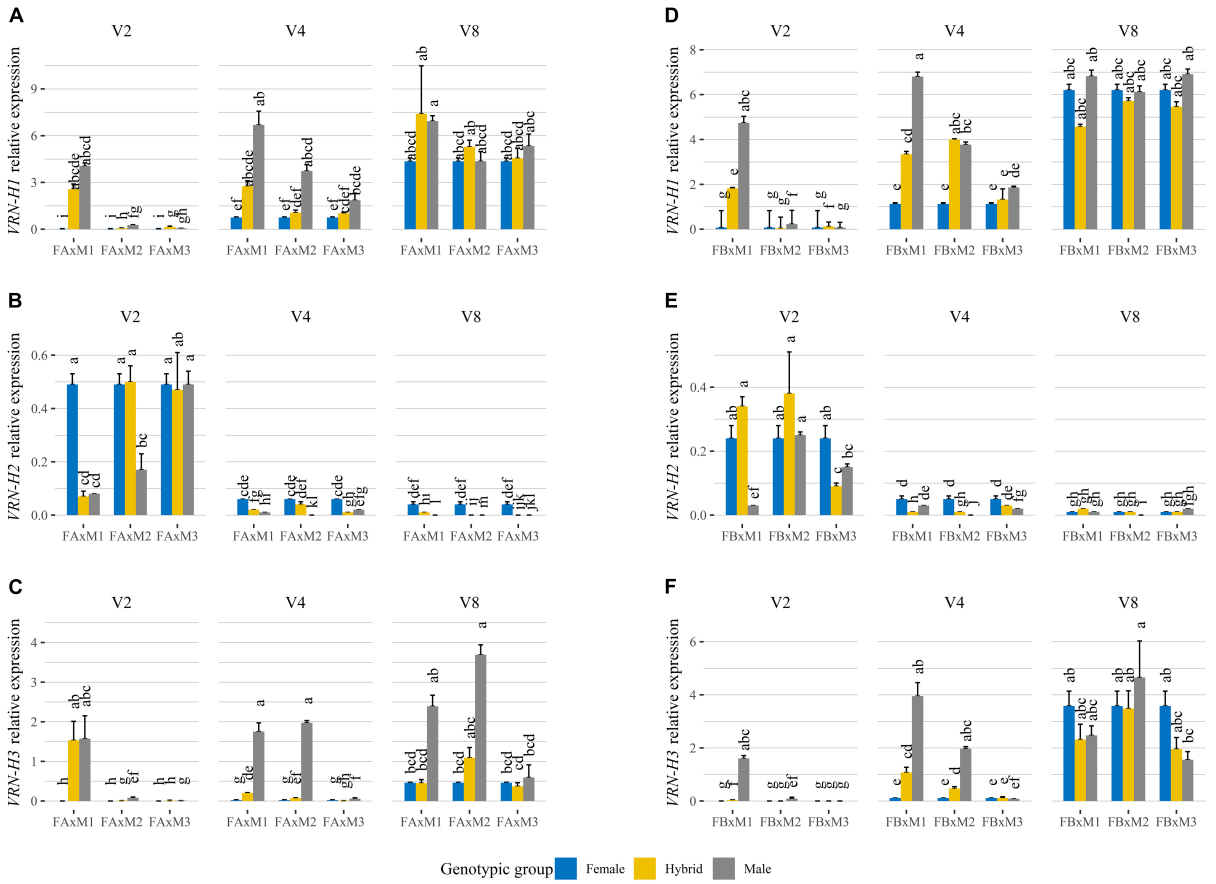


FIGURE 4 | Relative expression levels, at 35 days of growth in each treatment, of *VRN-H1* **(A,D)**, *VRN-H2* **(B,E)**, and *VRN-H3* **(C,F)** assayed by qRT-PCR in triads of barley genotypes (Female, Hybrid, Male) grown under 16 h light, in response to different vernalization treatments (V2: 2 weeks of vernalization, V4: 4 weeks of vernalization, V8: 8 weeks of vernalization; 4–8°C, 8 h light). Plots **(A–C)** correspond to Female A crosses (Batch A). Plots **(D–F)** correspond to Female B crosses (Batch B). Each plot is divided into three facets, each of them containing gene expression assayed for one vernalization treatment, and the three triads of genotypes composed of one female parent in blue, the male parent in gray, and the hybrid in yellow. The triads are represented as abbreviations of the crosses between the parents, e.g., FA × M1: Female A × Male 1. The results shown are normalized to the level of the housekeeping gene *Actin* for each genotype and treatment. Mean of 3 biological replicates. Error bars represent the SEM. For each gene and batch, bars with a different letter are significantly different at $P < 0.05$, according to ANOVA that included genotypes and all treatments.

the experiment carry the photoperiod-sensitive *PPD-H1* allele (**Supplementary Figure 4D**).

Expression of *PPD-H2* was detected in all genotypes that carried the gene, i.e., Male 3, Hybrid A3, and Hybrid B3 (**Supplementary Figures 4, 5**). *PPD-H2* expression was detected after 4 and 8 weeks of cold, but not in the 2-week vernalization treatment, confirming that a cold period is needed to induce its expression in winter genotypes. Expression in Male 3 and Hybrid B3 in the V4 and V8 treatments was similar, indicating the dominance of the active *PPD-H2* allele (**Supplementary Figure 5F**).

Associations Between Developmental Phases and Flowering Time Genes Expression

We performed a multiple factorial analysis (MFA) to examine patterns of relationships between developmental phases and gene expression averaged over the two sampling dates (**Supplementary Figure 6**). The expression of flowering inducers *VRN-H1*, *PPD-H1* (ns), *VRN-H3*, and *PPD-H2* showed a negative correlation with the length of developmental phases (**Figure 5** and **Supplementary Figure 1**), i.e., higher expression of gene inducers was related to earliness. On the contrary, the expression of flowering repressors *VRN-H2* and *HvODDSOC2* showed a positive correlation with the length of developmental phases and negative correlations with the flowering inducers, positioned in the opposite semicircle (**Figure 5**). The first dimension, accounting for almost 60% of the variance, summarized the time to reach the jointing stage and awn tipping, and the expression of the regulators involved in the control of the length of these periods, i.e., vernalization genes *VRN-H1* and *VRN-H2*, and *HvODDSOC2* and *VRN-H3*. The closeness of Z49 and Z31 vectors is explained by a correlation coefficient of 0.98 between the two variables. The duration of the lag Z31–Z49 had large loadings on the two dimensions and was relatively independent

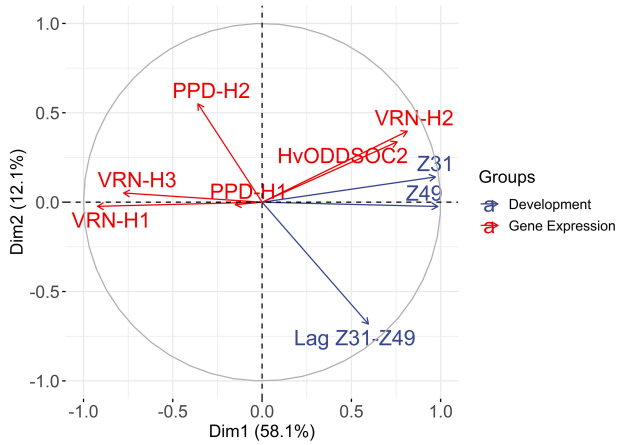


FIGURE 5 | Multiple factorial analysis (MFA), variable correlation circle. The plot shows the correlation of the quantitative variables with the MFA axes. Variables related to developmental phases are depicted in blue and variables related to gene expression are depicted in red. The expression of flowering promoters *VRN-H1*, *PPD-H1* (ns), *VRN-H3*, and *PPD-H2* shows a negative correlation with the length of developmental phases. On the contrary, the expression of flowering repressors *VRN-H2* and *HvODDSOC2* shows a positive correlation with the length of developmental phases and negative correlations with the promoters, positioned in the opposite semicircle. The first dimension represents mainly the time to reach the jointing stage and the genes that influence it, whereas the second dimension is more related to the late reproductive phase.

of the duration until Z31. The second dimension (12% of the variance) was related to the duration of the lag Z31-Z49, the expression of *PPD-H2,* and, to a lesser extent, *VRN-H2* and *HvODDSOC2*, reflecting the negative correlation between *PPD-H2* expression and the duration of the lag Z31–Z49. The variable *PPD-H1* was located close to the origin, indicating a poor representation on the factor map.

When plotting the individuals in the MFA, we could observe that the first axis mainly opposed the genotypes in the V8 and V2 treatments (**Supplementary Figure 7**). The genotypes Female A and Hybrid A3, carrying winter alleles at *VRN-H1*, showed the highest positive coordinates in the x-axis, which were positively correlated with a longer Z31 phase and a lower expression of *VRN-H1*. In contrast, genotypes characterized by a lower vernalization sensitivity, Male 1 and Male 2, showed the lowest negative coordinates, indicating shorter periods until the first node appearance and higher expression of *VRN-H1*. The second axis was essentially associated with genotypes Hybrid A2, Hybrid B3, Male 2, and Male 3, characterized by a short lag Z31–Z49 phase and low values of *VRN-H2* expression. On the opposite side of this axis, genotypes Female B and Hybrid A1 showed a consistently long late reproductive phase.

We detected variation in the length of the late reproductive phase, and its response to vernalization, which seemed related to the *VRN-H3* allele. We observed that those genotypes carrying the *vrn-H3c(1)* allele presented a short late reproductive phase independently of vernalization. Male 2 and its hybrids carry this allele and are represented by horizontal ellipses located on the positive side of the y-axis of the individuals' MFA for genotypes (**Supplementary Figure 8**). In contrast, those genotypes carrying one or two copies of the *vrn-H3d(1)* allele (Female A, Male 1, and derived hybrids) showed a constant and long duration of the late reproductive phase across treatments, and their ellipses are located in the negative coordinates of the y-axis (**Supplementary Figure 8**). Besides, Male 1 and Hybrid A1 were represented by small ellipses, pointing out a reduced variance in their responses across vernalization treatments. In contrast, Hybrid B3 showed a more vertical distribution, indicating higher variance for the second axis, related to lag Z31–Z49 phase, *PPD-H2,* and *VRN-H2* expression (**Supplementary Figure 8**).

DISCUSSION

Climate change is challenging current agricultural practices, posing questions about the best combinations of genotype × environment × management options for the near future (Cooper et al., 2021). Sheehan and Bentley (2021) recently pointed out the need for greater flexibility in varietal flowering time to sustain United Kingdom wheat productivity, a view that can be easily extended to barley, and to other geographical areas. Other adaptation strategies that extend the catalog of possibilities include shifts in the sowing date. Recent studies suggest shifting toward earlier sowings to offset climate change impacts and increase cereal yields in the future scenario (Zheng et al., 2012; Hunt et al., 2019). Phenological adjustment of barley hybrids requires acquiring detailed knowledge of the functioning of major flowering time genes in heterozygosis. Our experiment supposes the first step in this direction. We exposed a set of hybrids and their parents to a range of vernalization conditions, which has provided some new insights for phenology management in hybrid barley breeding.

Flowering Time in Hybrids Is Intermediate Between Parents

The Mediterranean-adapted lines used as pollinators were earlier than the central European elite lines used as female parents, whereas the hybrids were intermediate. This was true for both Z31 (jointing) and Z49 (awn tipping) stages, across all vernalization treatments, indicating the presence of additive inheritance. Additivity was not complete, however, as hybrids reached the Z31 stage significantly later than the average of parental lines (1.33 days, p-value < 0.001). In contrast, hybrids headed significantly earlier than the parental average (1.21 days, p-value < 0.01), due to a shorter late reproductive phase of hybrids, compared to the parental average (2.55 days, p-value < 0.001). The addition of the two phases resulted in slight heterosis toward earliness for Z49, which is a common finding in cereals, as reported for wheat (Borghi et al., 1988; Barbosa-Neto et al., 1996; Ahmed et al., 2000; Corbellini et al., 2002; Dreisigacker et al., 2005; Longin et al., 2013; Zhao et al., 2014; Al-Ashkar et al., 2020), triticale (Oettler et al., 2001), and barley (Zali and Allard, 1976; Oury et al., 2000; Bernhard et al., 2017). However, we

revealed a distinct gene action at each developmental stage, with the prevalence of additivity in the foundation phase (up to jointing), and a trend toward dominance for earliness in the construction phase. This was not unexpected, as the different genetic control of the length of the preanthesis phenological phases in winter cereals is well supported by strong experimental evidence (Slafer and Rawson, 1994; Miralles and Richards, 2000; González et al., 2002; Gol et al., 2017; Ochagavía et al., 2018).

Vernalization Mostly Affects the Foundation Growth Period, Which Is Controlled by Allelic Variation at *VRN-H1*/*VRN-H2* Genes

The vernalization treatments reduced the duration of the time until Z31 (72% on average, comparing V8 with V2), and lag Z31–Z49 (23% on average). Therefore, the sensitivity to vernalization mostly affected the vegetative and early reproductive phases, largely in agreement with the literature (Flood and Halloran, 1984; Griffiths et al., 1985; Roberts et al., 1988; Slafer and Rawson, 1994; Whitechurch et al., 2007), although strong effects of vernalization on the duration of the construction phase have also been reported (González et al., 2002). We observed this last effect only for the females.

Despite testing very different genotypes, the agreement between phenological development and gene expression supported our assumption that the range of responses to vernalization can largely be traced to the effect of the alleles present in *VRN-H1*. The multifactorial analysis indicated that the length until the reproductive transition (Z31) was associated with the pattern of expression of *VRN-H1* and *VRN-H2* genes, whose epistatic interaction controls the response to vernalization (von Zitzewitz et al., 2005). A novel finding of this study is that the length of the cold treatment needed to induce the expression of *VRN-H1*, as well as the degree of promotion toward flowering, depended on the allele constitution and dosage at *VRN-H1*. The dynamics of expression of *VRN-H1* alleles in the parents responded to the expectations of the gradual vernalization requirements induced by the three alleles. Thus, parents carrying the *VRN-H1-4* allele showed higher *VRN-H1* expression and accelerated development after just 2 weeks of cold; *VRN-H1-6* parents needed at least 4 weeks to reach the same point, whereas those carrying the *vrn-H1* allele required 8 weeks. The comparisons between homozygotes (parents and hybrids) indicated gradually decreasing vernalization requirements induced by alleles *vrn-H1*, *VRN-H1-6,* and *VRN-H1-4*, which confirms the gradation in the strength of flowering promotion displayed by the allelic series at *VRN-H1* (Takahashi and Yasuda, 1971; Szűcs et al., 2007; Casao et al., 2011a).

Additive Inheritance of *VRN-H1* Winter Alleles

The use of parental lines with different *VRN-H1* alleles provided the opportunity to assess the gene action at this locus. When *VRN-H1* alleles were confronted in the crosses, we observed intermediate Z31 and Z49 phenotypes (and *VRN-H1* expression) between the early and late alleles indicating additivity of the effect of winter *VRN-H1* alleles. In triads where there was no variation for *VRN-H1*, the phenotypic differences between genotypes were small, regardless of the vernalization treatment, and most likely due to other genes.

The prevalent view among geneticists indicates the dominance of the spring growth habit over the winter type (Takahashi and Yasuda, 1971; Dubcovsky et al., 2005; Fu et al., 2005). This view is supported by a dominant inheritance of the gene *VRN-H1* at the expression level in spring × winter crosses (Haas et al., 2020). Our results challenge this view. In fact, a review of the literature finds other results in agreement with ours. Some studies found hybrids with intermediate flowering dates between parents carrying spring *VRN-H1* and winter *vrn-H1* alleles. While complete dominance may occur in particular environmental conditions, experiments covering a wider and more realistic range of conditions revealed that additivity is the rule more than the exception in the vernalization process (Kóti et al., 2006; Szűcs et al., 2007). The additivity in the inheritance of winter *VRN-H1* alleles has agronomic implications. It expands the range of barley flowering time and vernalization responses available using hybrid combinations and can be used by breeders to fine-tune varietal vernalization needs to the target environments.

A reduced vernalization requirement, matching winter harshness level, may cause timely flowering and enhance yield. In fact, earliness conferred by *VRN-H1-4* was associated with increased grain yield in warm sites prone to the occurrence of terminal water stress (Mansour et al., 2014). Therefore, it seems a good choice to deploy in barley breeding for future scenarios in which current vernalization potential will be reduced, either in homozygosis or in hybrid combinations with other alleles.

The Construction Growth Period Shows a Dominant Inheritance Controlled by *FT*-Family Genes

Effects of *VRN-H1*, *VRN-H2*, and *PPD-H2* have been associated mainly with the length of the vegetative and early reproductive phases (Gol et al., 2017; Mulki et al., 2018). *VRN-H3* and *PPD-H1*, however, seem to affect the length of the late reproductive phase (Alqudah et al., 2014). In this experiment, *PPD-H2* expression apparently affected the duration of the jointing phase, but mostly the late reproductive phase, as also noticed by Casas et al. (2011).

We observed that the presence of *PPD-H2* modulated the responses of the *VRN-H1* alleles. In the two genotype triads involving a functional *PPD-H2* allele (including Male 3), we detected differences in the duration of development until the initiation of the jointing phase, which did not match the expectations based solely on their *VRN-H1* alleles. Male 3 and its hybrids showed a steeper reduction in development time, in response to moderate and complete vernalization than the female parents (both carrying a non-functional *ppd-H2* allele). However, this effect was absent in the low vernalization treatment: Female

B (*VRN-H1-6*) was earlier at Z49 than Male 3 (*vrn-H1*) at V2, but this order was reversed at V4. Concurrent with this crossover of cycle duration, we detected *PPD-H2* expression, in the male or the hybrid, only after 4 and 8 weeks of cold, but not in the 2-week vernalization treatment. This suggests that *PPD-H2* responds not only to photoperiod, but also to vernalization, and helps to accelerate development only after some vernalization has occurred (between 2 and 4 weeks in this case). This result agrees with Monteagudo et al. (2019b) who found that *PPD-H2* expression required some developmental trigger (either a cold period or advanced plant age) in winter barleys. Moreover, the earliness effect of *PPD-H2* was highly conspicuous in the shortening of the late reproductive phase of the hybrid and male when fully vernalized. As a result, hybrids carrying *PPD-H2* flowered at least as early as the earlier parent when vernalization was fully satisfied, indicating the dominance of the *PPD-H2* functional allele.

Boosting the development by ensuring timely completion of the cycle, *PPD-H2* is predominant in spring barleys. However, its agronomic merit in winter barley is not clear. Previous studies indicate that this gene acts as a safeguard mechanism in winter barleys, promoting spikelet initiation under short days, and reducing vernalization requirement under long days (Casao et al., 2011b; Mulki et al., 2018). Also, *PPD-H2* seems to have an adaptive role, confirmed by its influence on key agronomic traits (Cuesta-Marcos et al., 2009; Mansour et al., 2018; Sharma et al., 2018; Monteagudo et al., 2019a). We have shown that one single functional *PPD-H2* allele in a winter barley hybrid does accelerate flowering under insufficient vernalization (provided a minimum vernalization threshold is supplied). Therefore, our data support the role of *PPD-H2* as a source of earliness, to promote timely growth in warm winters with incomplete vernalization, which can be included in the formulation of hybrids for areas with mild winters.

On the other hand, *VRN-H3* allelic variation also contributed to differences in the length of the late reproductive phase. The *vrn-H3c(1)* allele was associated with a short late reproductive phase independently of vernalization, showing partial dominance in the hybrids. This allele combines an early promoter with an early intron haplotype and has been associated with the earliest flowering in both, a landrace collection of predominantly winter barleys (SBCC) (Casas et al., 2011), and a cross of two spring cultivars (Casas et al., 2021). In contrast, the *vrn-H3d(1)* allele, characterized by a late promoter and early intron haplotype, seems to confer a long and constant duration of the reproductive phase, independently of vernalization. The late reproductive phase determines the potential number of grains; therefore, this allele could be a stable resource for breeders aiming at high-yielding varieties.

Suggestions and Perspectives

This work shows the wide range of vernalization responses and flowering times that barley, and in particular hybrids, can display. From this information, the breeder can choose the earliness combination that best suits the conditions of each target environment. Here are some suggestions for allelic combinations that might work well, depending on the environment.

In climates where winters are long and cold, similar to our V8 treatment, and where we could expect that the vernalization needs will be completely satisfied, any of the tested hybrids would reach heading in a suitable date. However, to maximize yield, the breeder should choose a hybrid that matches the length of the LRP with the resource availability. In this sense, in environments prone to terminal stress, hybrids with a short LRP would have a better chance of escaping stress, which could be achieved with the *vrn-H3c(1)* allele or the *PPD-H2* allele. In contrast, in those environments where the end of the cycle benefits from optimal conditions, hybrids with a long LRP could enhance yield. In this scenario, the *vrn-H3d(1)* or *vrn-H3a(1)* allele could provide the effect we are looking for.

In those climates where winters are becoming warmer, comparable to our V4 treatment, where the cold needs might be compromised, the hybrid options that would flower on time are reduced. Hybrids carrying a dominant *PPD-H2* allele would be a suitable option when at least 4 weeks of cold were ensured. However, if the vernalization period decreased under 4 weeks, the delay in flowering time could negatively affect yield. A similar situation could be expected in hybrids with the *VRN-H1-6/vrn-H1* allelic combination. These could stand a certain decrease in the cold period duration, but not less than 4 weeks.

Under those circumstances, the safest option to ensure prompt flowering in the climate change scenario would be to use a hybrid with at least one *VRN-H1-4* allele or two *VRN-H1-6* alleles. Using these hybrid combinations should guarantee to maintain a suitable flowering even when the cold period is reduced under 4 weeks (V2 treatment).

In conclusion, although based on a small set of genotypes, we have demonstrated that hybrids can show a more nuanced response to insufficient vernalization than inbred lines. We also show that these phenotypic responses agree with the expression levels of main developmental genes. New options are proposed to manage time to flowering based on specific alleles and, particularly, the duration of developmental phases that build yield potential in hybrid barley. Our results highlight that hybrid combinations extend the available catalog of genetic responses to vernalization, which would be useful for adaptation to environmental conditions representing expected climate change trends.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in the **Supplementary Tables 6–9** of the **Supplementary Material** section.

AUTHOR CONTRIBUTIONS

EI and AC conceived this work. FC and EI obtained the plant material. MF-C carried out the phenotyping and data collection. AC and MF-C performed the laboratory work. EI and MF-C performed the statistical analyses. MF-C, AC, and

EI drafted the manuscript. All the authors read and approved the manuscript.

FUNDING

This research was supported by the contract "Iberia region hybrid barley variety development and understanding effects of adaptation genes in hybrids," between CSIC and Syngenta (Basel, Switzerland) Crop Protection AG, which included funding for MF-C scholarship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.827701/full#supplementary-material

REFERENCES

Ahmed, T. A., Tsujimoto, H., and Sasakuma, T. (2000). Identification of RFLP markers linked with heading date and its heterosis in hexaploid wheat. *Euphytica* 116, 111–119. doi: 10.1023/A:1004076826528

Al-Ashkar, I., Alotaibi, M., Refay, Y., Ghazy, A., Zakri, A., and Al-Doss, A. (2020). Selection criteria for high-yielding and early-flowering bread wheat hybrids under heat stress. *PLoS One* 15:e0236351. doi: 10.1371/journal.pone.0236351

Alqudah, A. M., Sharma, R., Pasam, R. K., Graner, A., Kilian, B., and Schnurbusch, T. (2014). Genetic Dissection of Photoperiod Response Based on GWAS of Pre-Anthesis Phase Duration in Spring Barley. *PLoS One* 9:e113120. doi: 10.1371/journal.pone.0113120

Barbosa-Neto, J. F., Sorrells, M. E., and Cisar, G. (1996). Prediction of heterosis in wheat using coefficient of parentage and RFLP-based estimates of genetic relationship. *Genome* 39, 1142–1149. g96-144 doi: 10.1139/

Bernhard, T., Friedt, W., Voss-Fels, K. P., Frisch, M., Snowdon, R. J., and Wittkop, B. (2017). Heterosis for biomass and grain yield facilitates breeding of productive dual-purpose winter barley hybrids. *Crop Sci.* 57, 2405–2418. doi: 10.2135/cropsci2016.10.0872

Bolaños, J., and Edmeades, G. O. (1993). Eight cycles of selection for drought tolerance in lowland tropical maize. II. Responses in reproductive behavior. *F. Crop. Res.* 31, 253–268. doi: 10.1016/0378-4290(93)90065-U

Borghi, B., Perenzin, M., and Nasht, R. J. (1988). Agronomic and qualitative characteristics of ten bread wheat hybrids produced using a chemical hybridizing agent. *Euphytica* 39, 185–194. BF00039872 doi: 10.1007/

Campoli, C., Shtaya, M., Davis, S. J., and von Korff, M. (2012). Expression conservation within the circadian clock of a monocot: natural variation at barley Ppd-H1 affects circadian expression of flowering time genes, but not clock orthologs. *BMC Plant Biol.* 12:97. doi: 10.1186/1471-2229-12-97

Campoli, C., and von Korff, M. (2014). "Genetic Control of Reproductive Development in Temperate Cereals," in *Advances in Botanical Research*, ed. F. Fornara (Cambridge: Academic Press), 131–158. doi: 10.1016/B978-0-12-417162-6.00005-5

Casao, M. C., Igartua, E., Karsai, I., Lasa, J. M., Gracia, M. P., and Casas, A. M. (2011a). Expression analysis of vernalization and day-length response genes in barley (Hordeum vulgare L.) indicates that VRNH2 is a repressor of PPDH2 (HvFT3) under long days. *J. Exp. Bot.* 62, 1939–1949. doi: 10.1093/jxb/erq382

Casao, M. C., Karsai, I., Igartua, E., Gracia, M. P., Veisz, O., and Casas, A. M. (2011b). Adaptation of barley to mild winters: A role for PPDH2. *BMC Plant Biol.* 11:164. doi: 10.1186/1471-2229-11-164

Casas, A. M., Djemel, A., Ciudad, F. J., Yahiaoui, S., Ponce, L. J., Contreras-Moreira, B., et al. (2011). HvFT1 (VrnH3) drives latitudinal adaptation in Spanish barleys. *Theor. Appl. Genet.* 122, 1293–1304. doi: 10.1007/s00122-011-1531-x

Casas, A. M., Gazulla, C. R., Monteagudo, A., Cantalapiedra, C. P., Moralejo, M., Pilar Gracia, M., et al. (2021). Candidate genes underlying QTL for flowering time and their interactions in a wide spring barley (Hordeum vulgare L.) cross. *Crop. J.* 9, 862–872. doi: 10.1016/J.CJ.2020.07.008

Cockram, J., Jones, H., Leigh, F. J., O'Sullivan, D., Powell, W., Laurie, D. A., et al. (2007). Control of flowering time in temperate cereals: Genes, domestication, and sustainable productivity. *J. Exp. Bot.* 58, 1231–1244. doi: 10.1093/jxb/erm042

Contreras-Moreira, B., Serrano-Notivoli, R., Mohammed, N. E., Cantalapiedra, C. P., Beguería, S., Casas, A. M., et al. (2019). Genetic association with high-resolution climate data reveals selection footprints in the genomes of barley landraces across the Iberian Peninsula. *Mol. Ecol.* 28, 1994–2012. doi: 10.1111/mec.15009

Cooper, M., Voss-Fels, K. P., Messina, C. D., Tang, T., and Hammer, G. L. (2021). Tackling G × E × M interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by predicting contributions of genetics and management to crop productivity. *Theor. Appl. Genet.* 134, 1625–1644. doi: 10.1007/s00122-021-03812-3

Corbellini, M., Perenzin, M., Accerbi, M., Vaccino, P., and Borghi, B. (2002). Genetic diversity in bread wheat, as revealed by coefficient of parentage and molecular markers, and its relationship to hybrid performance. *Euphytica* 123, 273–285. doi: 10.1023/A:1014946018765

Craufurd, P. Q., and Wheeler, T. R. (2009). Climate change and the flowering time of annual crops. *J. Exp. Bot.* 60, 2529–2539. doi: 10.1093/jxb/erp196

Cuesta-Marcos, A., Casas, A. M., Hayes, P. M., Gracia, M. P., Lasa, J. M., Ciudad, F., et al. (2009). Yield QTL affected by heading date in Mediterranean grown barley. *Plant Breed.* 128, 46–53. doi: 10.1111/j.1439-0523.2008.01510.x

Deng, W., Casao, M. C., Wang, P., Sato, K., Hayes, P. M., Finnegan, E. J., et al. (2015). Direct links between the vernalization response and other key traits of cereal crops. *Nat. Commun.* 6, 5882. doi: 10.1038/ncomms6882

Distelfeld, A., Li, C., and Dubcovsky, J. (2009). Regulation of flowering in temperate cereals. *Curr. Opin. Plant Biol.* 12, 178–184. doi: 10.1016/j.pbi.2008.12.010

Dreisigacker, S., Melchinger, A. E., Zhang, P., Ammar, K., Flachenecker, C., Hoisington, D., et al. (2005). Hybrid performance and heterosis in spring bread wheat, and their relations to SSR-based genetic distances and coefficients of parentage. *Euphytica* 144, 51–59. doi: 10.1007/s10681-005-4053-2

Dubcovsky, J., Chen, C., and Yan, L. (2005). Molecular characterization of the allelic variation at the VRN-H2 vernalization locus in barley. *Mol. Breed.* 15, 395–407. doi: 10.1007/s11032-005-0084-6

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *J. Stat. Softw.* 48, 1–18. doi: 10.18637/jss.v048.i04

Evans, L. (1996). *Crop Evolution, Adaptation and Yield*. (Cambridge: Cambridge university press.)

Fatima, Z., Ahmed, M., Hussain, M., Abbas, G., Ul-Allah, S., Ahmad, S., et al. (2020). The fingerprints of climate warming on cereal crops phenology and adaptation options. *Sci. Rep.* 10:18013. doi: 10.1038/s41598-020-74740-3

Faure, S., Higgins, J., Turner, A., and Laurie, D. A. (2007). The *FLOWERING LOCUS T* -Like Gene Family in Barley (*Hordeum vulgare*). *Genetics* 176, 599–609. doi: 10.1534/genetics.106.069500

Fernández-Calleja, M., Casas, A. M., and Igartua, E. (2021). Major flowering time genes of barley: allelic diversity, effects, and comparison with wheat. *Theor. Appl. Genet.* 134, 1867–1897. doi: 10.1007/s00122-021-03824-z

Flohr, B. M., Hunt, J. R., Kirkegaard, J. A., Evans, J. R., Trevaskis, B., Zwart, A., et al. (2018). Fast winter wheat phenology can stabilise flowering date and maximise grain yield in semi-arid Mediterranean and temperate environments. *F. Crop. Res.* 223, 12–25. doi: 10.1016/j.fcr.2018.03.021

Flood, R. G., and Halloran, G. M. (1984). Basic Development Rate in Spring Wheat. *Agron. J.* 76, 260–264. doi: 10.2134/agronj1984.00021962007600020021x

Fu, D., Szűcs, P., Yan, L., Helguera, M., Skinner, J. S., Von Zitzewitz, J., et al. (2005). Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Mol. Genet. Genomics* 273, 54–65. doi: 10.1007/s00438-004-1095-4

Gol, L., Tomé, F., and von Korff, M. (2017). Floral transitions in wheat and barley: interactions between photoperiod, abiotic stresses, and nutrient status. *J. Exp. Bot.* 68, 1399–1410. doi: 10.1093/jxb/erx055

González, A., Martín, I., and Ayerbe, L. (1999). Barley yield in water-stress conditions. The influence of precocity, osmotic adjustment and stomatal conductance. *F. Crop. Res.* 62, 23–34. doi: 10.1016/S0378-4290(99)00002-7

González, F. G., Slafer, G. A., and Miralles, D. J. (2002). Vernalization and photoperiod responses in wheat pre-flowering reproductive phases. *F. Crop. Res.* 74, 183–195. doi: 10.1016/S0378-4290(01)00210-6

Gouache, D., Bogard, M., Pegard, M., Thepot, S., Garcia, C., Hourcade, D., et al. (2017). Bridging the gap between ideotype and genotype: Challenges and prospects for modelling as exemplified by the case of adapting wheat (Triticum aestivum L.) phenology to climate change in France. *F. Crop. Res.* 202, 108–121. doi: 10.1016/j.fcr.2015.12.012

Gracia, M. P., Mansour, E., Casas, A. M., Lasa, J. M., Medina, B., Molina-Cano, J. L., et al. (2012). Progress in the Spanish National Barley Breeding Program. *Spanish J. Agric. Res.* 10, 741–751. doi: 10.5424/sjar/2012103-2613

Greenup, A. G., Sasani, S., Oliver, S. N., Talbot, M. J., Dennis, E. S., Hemming, M. N., et al. (2010). ODDSOC2 is a MADS box floral repressor that is down-regulated by vernalization in temperate cereals. *Plant Physiol.* 153, 1062–1073. doi: 10.1104/pp.109.152488

Griffiths, F. E. W., Lyndon, R. F., and Bennett, M. D. (1985). The Effects of Vernalization on the Growth of the Wheat Shoot Apex. *Ann. Bot.* 56, 501–511. doi: 10.1093/oxfordjournals.aob.a087035

Guerra, D., Morcia, C., Badeck, F., Rizza, F., Delbono, S., Francia, E., et al. (2021). Extensive allele mining discovers novel genetic diversity in the loci controlling frost tolerance in barley. *Theor. Appl. Genet.* [Epub Online ahead of print]. doi: 10.1007/S00122-021-03985-X

Haas, M., Himmelbach, A., and Mascher, M. (2020). The contribution of cis- and trans-acting variants to gene regulation in wild and domesticated barley under cold stress and control conditions. *J. Exp. Bot.* 71, 2573–2584. doi: 10.1093/jxb/eraa036

Hemming, M. N., Fieg, S., James Peacock, W., Dennis, E. S., and Trevaskis, B. (2009). Regions associated with repression of the barley (Hordeum vulgare) VERNALIZATION1 gene are not required for cold induction. *Mol. Genet. Genomics* 282, 107–117. doi: 10.1007/s00438-009-0449-3

Hemming, M. N., Peacock, W. J., Dennis, E. S., and Trevaskis, B. (2008). Low-Temperature and Daylength Cues Are Integrated to Regulate FLOWERING LOCUS T in Barley. *Plant Physiol.* 147, 355–366. doi: 10.1104/pp.108.116418

Hunt, J. R., Lilley, J. M., Trevaskis, B., Flohr, B. M., Peake, A., Fletcher, A., et al. (2019). Early sowing systems can boost Australian wheat yields despite recent climate change. *Nat. Clim. Chang.* 9, 244–247. doi: 10.1038/s41558-019-0417-9

Karsai, I., Szűcs, P., Mészáros, K., Filichkina, T., Hayes, P. M., Skinner, J. S., et al. (2005). The Vrn-H2 locus is a major determinant of flowering time in a facultative × winter growth habit barley (Hordeum vulgare L.) mapping population. *Theor. Appl. Genet.* 110, 1458–1466. doi: 10.1007/s00122-005-1979-7

Kassambara, A., and Mundt, F. (2017). *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. 337–354. Available online at: https://github.com/kassambara/factoextra/issues (Accessed on Dec 11, 2020)

Kikuchi, R., Kawahigashi, H., Ando, T., Tonooka, T., and Handa, H. (2009). Molecular and Functional Characterization of PEBP Genes in Barley Reveal the Diversification of Their Roles in Flowering. *Plant Physiol.* 149, 1341–1353. doi: 10.1104/pp.108.132134

Kóti, K., Karsai, I., Sz, P., Horváth, C., Mészáros, K., Kiss, G. B., et al. (2006). Validation of the two-gene epistatic model for vernalization response in a winter × spring barley cross. *Euphytica* 152, 17–24. doi: 10.1007/s10681-006-9170-z

Laurie, D. A., Pratchett, N., Snape, J. W., and Bezant, J. H. (1995). RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter × spring barley (*Hordeum vulgare* L.) cross. *Genome* 38, 575–585. doi: 10.1139/g95-074

Lê, S., Josse, J., Rennes, A., and Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *JSS J. Stat. Softw.* 25, 1–18. doi: 10.18637/jss.v025.i01

Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R package version* .

Longin, C. F. H., Gowda, M., Mühleisen, J., Ebmeyer, E., Kazman, E., Schachschneider, R., et al. (2013). Hybrid wheat: Quantitative genetic parameters and consequences for the design of breeding programs. *Theor. Appl. Genet.* 126, 2791–2801. doi: 10.1007/s00122-013-2172-z

Longin, C. F. H., Mühleisen, J., Maurer, H., Zhang, H., Gowda, M., and Reif, J. C. (2012). Hybrid breeding in autogamous cereals. *Theor. Appl. Genet.* 125, 1087–1096. doi: 10.1007/s00122-012-1967-7

Loscos, J., Igartua, E., Contreras-Moreir, B., Pilar Gracia, M., and Casas, A. M. (2014). HvFT1 polymorphism and effect—survey of barley germplasm and expression analysis. *Front. Plant Sci.* 5:251. doi: 10.3389/fpls.2014.00251

Mansour, E., Casas, A. M., Gracia, M. P., Molina-Cano, J. L., Moralejo, M., Cattivelli, L., et al. (2014). Quantitative trait loci for agronomic traits in an elite barley population for Mediterranean conditions. *Mol. Breed.* 33, 249–265. doi: 10.1007/s11032-013-9946-5

Mansour, E., Moustafa, E. S. A., Qabil, N., Abdelsalam, A., Wafa, H. A., Kenawy, A., et al. (2018). Assessing different barley growth habits under Egyptian conditions for enhancing resilience to climate change. *F. Crop. Res.* 224, 67–75. doi: 10.1016/j.fcr.2018.04.016

Miralles, D. J., and Richards, R. A. (2000). Responses of leaf and tiller emergence and primordium initiation in wheat and barley to interchanged photoperiod. *Ann. Bot.* 85, 655–663. doi: 10.1006/anbo.2000.1121

Monteagudo, A., Casas, A. M., Cantalapiedra, C. P., Contreras-Moreira, B., Gracia, M. P., and Igartua, E. (2019a). Harnessing novel diversity from landraces to improve an elite barley variety. *Front. Plant Sci.* 10:434. doi: 10.3389/fpls.2019.00434

Monteagudo, A., Igartua, E., Contreras-Moreira, B., Gracia, M. P., Ramos, J., Karsai, I., et al. (2019b). Fine-tuning of the flowering time control in winter barley: The importance of HvOS2 and HvVRN2 in non-inductive conditions. *BMC Plant Biol.* 19:113. doi: 10.1186/s12870-019-1727-9

Mühleisen, J., Maurer, H. P., Stiewe, G., Bury, P., and Reif, J. C. (2013). Hybrid Breeding in Barley. *Crop. Sci.* 53, 819–824. doi: 10.2135/CROPSCI2012.07.0411

Mühleisen, J., Piepho, H.-P., Maurer, H. P., Longin, C. F. H., and Reif, J. C. (2014). Yield stability of hybrids versus lines in wheat, barley, and triticale. *Theor. Appl. Genet.* 127, 309–316. doi: 10.1007/s00122-013-2219-1

Mulki, M. A., Bi, X., and von Korff, M. (2018). FLOWERING LOCUS T3 Controls Spikelet Initiation But Not Floral Development. *Plant Physiol.* 178, 1170–1186. doi: 10.1104/pp.18.00236

Mulki, M. A., and von Korff, M. (2016). CONSTANS Controls Floral Repression by Up-Regulating VERNALIZATION2 (VRN-H2) in Barley. *Plant Physiol.* 170, 325–337. doi: 10.1104/pp.15.01350

Nitcher, R., Distelfeld, A., Tan, C., Yan, L., and Dubcovsky, J. (2013). Increased copy number at the HvFT1 locus is associated with accelerated flowering time in barley. *Mol. Genet. Genomics* 288, 261–275. doi: 10.1007/s00438-013-0746-8

Ochagavía, H., Prieto, P., Savin, R., Griffiths, S., and Slafer, G. A. (2018). Dynamics of leaf and spikelet primordia initiation in wheat as affected by Ppd-1a alleles under field conditions. *J. Exp. Bot.* 69, 2621–2631. doi: 10.1093/jxb/ery104

Oettler, Becker, and Hoppe. (2001). Heterosis for yield and other agronomic traits of winter triticale F1 and F2 hybrids. *Plant Breed.* 120, 351–353. doi: 10.1046/j.1439-0523.2001.00624.x

Olesen, J. E., Trnka, M., Kersebaum, K. C., Skjelvåg, A. O., Seguin, B., Peltonen-Sainio, P., et al. (2010). Impacts and adaptation of European crop production systems to climate change. *Eur. J. Agron.* 34, 96–112. doi: 10.1016/j.eja.2010.11.003

Oliver, S. N., Deng, W., Casao, M. C., and Trevaskis, B. (2013). Low temperatures induce rapid changes in chromatin state and transcript levels of the cereal VERNALIZATION1 gene. *J. Exp. Bot.* 64, 2413–2422. doi: 10.1093/jxb/ert095

Oury, F., Brabant, P., Bérard, P., and Pluchard, P. (2000). Predicting hybrid value in bread wheat: Biometric modelling based on a" top-cross" design. *Theor. Appl. Genet.* 100, 96–104. doi: 10.1007/PL00002905

Pagès, J. (2002). Analyse factorielle multiple appliquée aux variables qualitatives et aux données mixtes. *Rev. Stat. appliquée* 50, 5–37.

Porter, J. R., Xie, L., Challinor, A. J., Chhetri Usa, N., Garrett, K., Aggarwal, P., et al. (2014). "Food security and food production systems," in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, et al. (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press), 485–553.

R Core Team. (2013). *R: A language and environment for statistical computing*.(Vienna, Aus: R Foundation for Statistical Computing)

Roberts, E. H., Summerfield, R. J., Cooper, J. P., and Ellis, R. H. (1988). Environmental Control of Flowering in Barley (Hordeum vulgare L.). I. Photoperiod Limits to Long-day Responses, Photoperiod-insensitive Phases and Effects of Low-temperature and Short-day Vernalization. *Ann. Bot.* 62, 127–144. doi: 10.1093/oxfordjournals.aob.a087644

Ruelens, P., De Maagd, R. A., Proost, S., Theißen, G., Geuten, K., and Kaufmann, K. (2013). FLOWERING LOCUS C in monocots and the tandem origin of angiosperm-specific MADS-box genes. *Nat. Commun.* 4:2280. doi: 10.1038/ncomms3280

Saadi, S., Todorovic, M., Tanasijevic, L., Pereira, L. S., Pizzigalli, C., and Lionello, P. (2015). Climate change and Mediterranean agriculture: Impacts on winter wheat and tomato crop evapotranspiration, irrigation requirements and yield. *Agric. Water Manag.* 147, 103–115. doi: 10.1016/j.agwat.2014.05.008

Sasani, S., Hemming, M. N., Oliver, S. N., Greenup, A., Tavakkol-Afshari, R., Mahfoozi, S., et al. (2009). The influence of vernalization and daylength on expression of flowering-time genes in the shoot apex and leaves of barley (*Hordeum vulgare*). *J. Exp. Bot.* 60, 2169–2178. doi: 10.1093/jxb/erp098

Sharma, R., Draicchio, F., Bull, H., Herzig, P., Maurer, A., Pillen, K., et al. (2018). Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding. *J. Exp. Bot.* 69, 3811–3822. doi: 10.1093/jxb/ery178

Sheehan, H., and Bentley, A. (2021). Changing times: Opportunities for altering winter wheat phenology. *Plants People Planet* 3, 113–123. doi: 10.1002/ppp3.10163

Slafer, G. A., and Rawson, H. M. (1994). Sensitivity of Wheat Phasic Development to Major Environmental Factors: a Re-Examination of Some Assumptions Made by Physiologists and Modellers. *Funct. Plant Biol.* 21, 393–426. doi: 10.1071/PP9940393

Stratonovitch, P., and Semenov, M. A. (2015). Heat tolerance around flowering in wheat identified as a key trait for increased yield potential in Europe under climate change. *J. Exp. Bot.* 66, 3599–3609. doi: 10.1093/jxb/erv070

Szűcs, P., Skinner, J. S., Karsai, I., Cuesta-Marcos, A., Haggard, K. G., Corey, A. E., et al. (2007). Validation of the VRN-H2/VRN-H1 epistatic model in barley reveals that intron length variation in VRN-H1 may account for a continuum of vernalization sensitivity. *Mol. Genet. Genomics* 277, 249–261. doi: 10.1007/s00438-006-0195-8

Takahashi, R., and Yasuda, S. (1971). "Genetics of earliness and growth habit in barley," in *Barley genetics II. Proceeding 2nd International Barley Genetics Symposium*, ed. R. A. Nilan (Pullman: Washington State University Press), 388–408.

Tao, F., Rötter, R. P., Palosuo, T., Díaz-Ambrona, C. G. H., Mínguez, M. I., Semenov, M. A., et al. (2017). Designing future barley ideotypes using a crop model ensemble. *Eur. J. Agron.* 82, 144–162. doi: 10.1016/j.eja.2016.10.012

Tondelli, A., Francia, E., Visioni, A., Comadran, J., Mastrangelo, A. M., Akar, T., et al. (2014). QTLs for barley yield adaptation to Mediterranean environments in the "Nure" x "Tremois" biparental population. *Euphytica* 197, 73–86. doi: 10.1007/s10681-013-1053-5

Trevaskis, B., Bagnall, D. J., Ellis, M. H., Peacock, W. J., and Dennis, E. S. (2003). MADS box genes control vernalization-induced flowering in cereals. *Proc. Natl. Acad. Sci.U.S.A.* 100, 13099–13104. doi: 10.1073/pnas.1635053100

Trevaskis, B., Hemming, M. N., Dennis, E. S., and Peacock, W. J. (2007). The molecular basis of vernalization-induced flowering in cereals. *Trends Plant Sci.* 12, 352–357. doi: 10.1016/j.tplants.2007.06.010

Trevaskis, B., Hemming, M. N., Peacock, W. J., and Dennis, E. S. (2006). HvVRN2 responds to daylength, whereas HvVRN1 is regulated by vernalization and developmental status. *Plant Physiol.* 140, 1397–1405. doi: 10.1104/pp.105.073486

Trnka, M., Rötter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Žalud, Z., et al. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. *Nat. Clim. Chang.* 4, 637–643. doi: 10.1038/nclimate2242

Turner, A., Beales, J., Faure, S., Dunford, R. P., and Laurie, D. A. (2005). The Pseudo-Response Regulator Ppd-H1 Provides Adaptation to Photoperiod in Barley. *Science* 310, 1031–1034. doi: 10.1126/science.1117619

von Zitzewitz, J., Szűcs, P., Dubcovsky, J., Yan, L., Francia, E., Pecchioni, N., et al. (2005). Molecular and Structural Characterization of Barley Vernalization Genes. *Plant Mol. Biol.* 59, 449–467. doi: 10.1007/s11103-005-0351-2

Whitechurch, E. M., Slafer, G. A., and Miralles, D. J. (2007). Variability in the duration of stem elongation in wheat genotypes and sensitivity to photoperiod and vernalization. *J. Agron. Crop. Sci.* 193, 131–137. doi: 10.1111/j.1439-037X.2007.00259.x

Wiegmann, M., Maurer, A., Pham, A., March, T. J., Al-Abdallat, A., Thomas, W. T. B., et al. (2019). Barley yield formation under abiotic stress depends on the interplay between flowering time genes and environmental cues. *Sci. Rep.* 9:6397. doi: 10.1038/s41598-019-42673-1

Yahiaoui, S., Igartua, E., Moralejo, M., Ramsay, L., Molina-Cano, J. L., Ciudad, F. J., et al. (2008). Patterns of genetic and eco-geographical diversity in Spanish barleys. *Theor. Appl. Genet.* 116, 271–282. doi: 10.1007/s00122-007-0665-3

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., et al. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci.U.S.A.* 103, 19581–19586. doi: 10.1073/pnas.0607142103

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., et al. (2004). The Wheat VRN2 Gene Is a Flowering Repressor Down-Regulated by Vernalization. *Science* 303, 1640–1644. doi: 10.1126/science.1094305

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene VRN1. *Proc. Natl. Acad. Sci.U.S.A.* 100, 6263–6268. doi: 10.1073/pnas.0937399100

Yang, C., Fraga, H., van Ieperen, W., Trindade, H., and Santos, J. A. (2019). Effects of climate change and adaptation options on winter wheat yield under rainfed Mediterranean conditions in southern Portugal. *Clim. Change* 154, 159–178. doi: 10.1007/s10584-019-02419-4

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zali, A. A., and Allard, R. W. (1976). The effect of level of heterozygosity on the performance of hybrids between isogenic lines of barley. *Genetics* 84, 765–775. doi: 10.1093/genetics/84.4.765

Zhao, Y., Mette, M. F., Gowda, M., Longin, C. F. H., and Reif, J. C. (2014). Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity* 112, 638–645. doi: 10.1038/hdy.2014.1

Zheng, B., Chenu, K., Fernanda Dreccer, M., and Chapman, S. C. (2012). Breeding for the future: What are the potential impacts of future frost and heat events on sowing and flowering time requirements for Australian bread wheat (Triticum aestivium) varieties? *Glob. Chang. Biol.* 18, 2899–2914. doi: 10.1111/j.1365-2486.2012.02724.x

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: 20 April 2022
doi: 10.3389/fpls.2022.882946

GhSOC1s Evolve to Respond Differently to the Environmental Cues and Promote Flowering in Partially Independent Ways

*Limei Ma and Yuanyuan Yan**

State Key Laboratory of North China Crop Improvement and Regulation, Key Laboratory for Crop Germplasm Resources of Hebei, College of Agronomy, Hebei Agricultural University, Baoding, China

Gossypium hirsutum is most broadly cultivated in the world due to its broader adaptation to the environment and successful breeding of early maturity varieties. However, how cotton responds to environmental cues to adjust flowering time to achieve reproductive success is largely unknown. *SOC1* functions as an essential integrator for the endogenous and exogenous signals to maximize reproduction. Thus we identified six *SOC1*-like genes in Gossypium that clustered into two groups. *GhSOC1-1* contained a large intron and clustered with monocot *SOC1s*, while *GhSOC1-2/3* were close to dicot *SOC1s*. *GhSOC1s* expression gradually increased during seedling development suggesting their conserved function in promoting flowering, which was supported by the early flowering phenotype of *35S:GhSOC1-1 Arabidopsis* lines and the delayed flowering of cotton silencing lines. Furthermore, *GhSOC1-1* responded to short-day and high temperature conditions, while *GhSOC1-2* responded to long-day conditions. *GhSOC1-3* might function to promote flowering in response to low temperature and cold. Taken together, our results demonstrate that *GhSOC1s* respond differently to light and temperature and act cooperatively to activate *GhLFY* expression to promote floral transition and enlighten us in cotton adaptation to environment that is helpful in improvement of cotton maturity.

Keywords: *Gossypium hirsutum* L., *SOC1*-like gene, flowering time control, environmental response, evolution

OPEN ACCESS

Edited by:
Yang Zhu,
Zhejiang University, China

Reviewed by:
Geng-Qing Huang,
Central China Normal University, China
Tianlun Zhao,
Zhejiang University, China

***Correspondence:**
Yuanyuan Yan
selina3001630016@163.com

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

Received: *24 February 2022*
Accepted: *23 March 2022*
Published: *20 April 2022*

Citation:
Ma L and Yan Y (2022) GhSOC1s Evolve to Respond Differently to the Environmental Cues and Promote Flowering in Partially Independent Ways. Front. Plant Sci. 13:882946. doi: 10.3389/fpls.2022.882946

INTRODUCTION

Cotton is an important cash crop that produce the natural textile fiber, which supports the life of an estimated 150 million people (Amrouk et al., 2021). The upland cotton (*Gossypium hirsutum* L. AADD, $2n = 52$) is broadestly cultivated to provide the world major cotton fibers. Cotton origins from subtropics. Consistent to the unnecessary adaption to winter, the main regulator in vernalization, *FLC* (*FLOWERING LOCUS C*), was absent in the *Gossypium* genomes (Nardeli et al., 2018). It evolved with a broader adaptation to the environment and the sensitiveness to short-day condition is lost during domestication (Hao et al., 2008; Ren et al., 2017), which largely extent planting area of cotton. Epigenome and GWAS analysis have identified genetic and

epigenetic changes contributing to loss sensitivity of upland cotton to short-day photoperiod (Song et al., 2017; Li et al., 2021b). However, the knowledge for cotton flowering time control is deficient. The question how cotton response to environment to initiate floral transition at the appropriate time remains unknown.

Floral transition divides plant growth into the vegetative and reproductive stages. Plant only grows leaves, stems and roots in the vegetative growth until floral transition when the reproductive organs initiate. In *Arabidopsis*, *CONSTANS* (*CO*) acts as the central regulator of photoperiod pathway. Overexpression of *CO* causes early flowering phenotype, which could be rescued by mutation in a gene named *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*). *SOC1* integrates the signals from photoperiod, age, temperature and GA to initiate the differentiation of flower bud, which is essential for plants to flowering at the appropriate time to achieve the maximal reproductive success (Lee and Lee, 2010). *SOC1* transcription can be activated by effectors of photoperiod and GA signaling (Yoo et al., 2005), and is suppressed by *SHORT VEGETATIVE PHASE* (*SVP*) that responses to the autonomous and GA pathway to avoid low propagation resulting from early flowering (Turck et al., 2008). The SOC1-SPL model governs the signals from photoperiod and GA to control flowering (Jung et al., 2012). During vernalization, *FLC* is fine tuned to directly inhibit *SOC1* expression, thus the plants can overcome the cold winter (Li et al., 2008). In the shoot apical meristem (SAM), SOC1 and AGL24 abundance is maintained by positivist feedback loop to ensure the entrance of SOC1 into nucleus with the help of AGL24, where the SOC1/AGL24 complex directly activates *LFY* transcription (Lee et al., 2008). Furthermore, SOC1 binds loci of a large number of flowering time regulators including the majority of its own repressors to establish the double-negative feedback loop (Immink et al., 2012). It is more likely that SOC1 situates in the center of the flowering regulatory network containing multiple regulatory and feedback loops that serves as a molecular switch of floral transition.

In *Arabidopsis*, *soc1* mutant flowers late, while overexpression of *SOC1* accelerates flowering (Borner et al., 2008; Lee and Lee, 2010). *SOC1*-like genes in other species also function as flowering time promoters, such as orchard *DoSOC1* (Ding et al., 2013) and peony *PsSOC1* (Wang et al., 2015). Ectopic expression of the K domain of blue berry *VcSOC1* gene also promotes flowering of tobacco, suggesting that the K-box is the functional domain of *VcSOC1* (Song et al., 2013). Although the function of *SOC1* homologs is conserved, there are evidence supporting functional divergence of *SOC1*-like genes. *FaSOC1* from cultivated strawberry promotes flowering, while *FvSOC1* from wild strawberry suppresses flowering to maintain the vegetative growth (Lei et al., 2013; Mouhu et al., 2013). Ectopic expression of the *SOC1* homolog from *Gerbra hybrida* cannot affect flowering time, but cause defects of floral organs (Ruokolainen et al., 2011). Functional and transcriptional analysis of *SOC1*-like genes in *Actinidia chinensis* suggest that they are involved in seed dormancy and may evolved to loss the function of flowering time control (Voogd et al., 2015). *SOC1* also responds to photoperiod, draft, cold and high temperature to regulate growth, stomatal opening and chloroplast biogenesis (Ryu et al., 2009; Richter et al., 2013; Kimura et al., 2015; Wang et al., 2019).

Considering the comprehensive response of *SOC1* to endogenous and exogenous signals, it is imperative to identify the function of *SOC1* homologs in Gossypium. In previous study, *GhSOC1* was cloned from CCRI36, which could promote flowering of *Arabidopsis* when overexpressed. But overexpression of *GhSOC1* in cotton only affected floral organ development rather than flowering time (Zhang et al., 2016). Therefore, the clarification of functions of different *SOC1*-like genes is vital for understanding the regulatory mechanism of cotton flowering time control. In this study, *SOC1*-like genes were identified in Gossypium genomes and their evolution were analyzed. Then the expression characters of different *GhSOC1s* were investigation and *GhSOC1-1* function was further investigated by overexpression and silencing. The results showed that cotton *SOC1*-like genes evolved divergently to respond differently to light and temperature and promote flowering in a cooperative way, which indicates genomic changes among GhSOC1 locus relating to adaptation. Our findings draw a dynamic regulatory model participated by homologies in tetraploid cotton, which help us to understand the mechanism of cotton flowering time control in response to the environment.

MATERIALS AND METHODS

Identification of *SOC1*-Like Genes in Cotton

The protein sequences of different versions in diploid and tetraploid cotton were downloaded from CottonGen.[1] SOC1 homologs were obtained through tBLSATp searches using AtSOC1 (AT2G45660) as query against the diploid and tetraploid genomes, respectively. The conserved MADS and K-box domains were further confirmed by NCBI CD-Search.[2] The identified cotton SOC1 homologs were listed in **Supplementary Table 1**.

Bioinformatics Analysis

To construct the phylogenetic tree, we aligned full-length protein sequences of all cotton *SOC1*-like genes together with the other eighteen species that were listed in **Supplementary Table 2** and used MEGA 7.0 to construct the phylogenetic tree (Neighbor join, Bootstrap 1,000 and 50% cutoff values) (Kumar et al., 2016). DNAMAN was applied for alignment.

The gene structure was analyzed on GSDS.[3] WebLogo3 was applied online[4] to build weblogo diagram. The three-dimensional structure of proteins was predicted on the website SWISS-MODEL.[5] The *cis*-elements were predicted on Plant CARE.[6] The CarG-box elements were searched by tbtools

[1]https://www.cottongen.org/

[2]https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

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

[4]http://weblogo.threeplusone.com/create.cgi

[5]https://swissmodel.expasy.org/

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

software and visualized according to the position. Transcriptome data of floral organs was downloaded from the NCBI SRA (Sequence Read Archive) database (Genome sequencing project accession: PRJNA248163). The FPKM of each *GhSOC1* gene was normalized to $\log_2$ (1+FPKM), and expression heat maps were drawn using TBtools software.

Plant Growth and Sample Collection

G. hirsutum L. TM-1 were grown in the green house (28°C, 16 h light/8 h dark). The roots, stems and true leaves were collected, respectively, when the first true leaf flattened, and the flowers and fibers were sampled at 0 and 10 days post anthesis, respectively. These samples were used for tissue specific expression analysis. The aerial parts of true leaves were collected when the third, fourth and fifth true leaf flattened, respectively, for temporal expression analysis. When the cotton seedlings grew to the third true leaf stage under long-day (LD), they were shifted to short-day (SD, 28°C), high temperature (35°C) and low temperature (18°C) conditions and the whole seedlings were collected after 48 h for quick response analysis. Upland cotton TM-1 was continuously cultivated under different day length (28°C) and temperatures (18, 22, 28, 35°C under LD) and the whole seedling were sampled at the third and fifth true-leaf stages for expression comparison.

Arabidopsis plants were grown under 23°C, 16 h light/8 h dark. The rosette leaves of the T1 transgenic lines were collected for DNA extraction followed by genotyping with gene specific primers. The seedlings of T3 homozygous lines were collected 9 days after germination for expression analysis.

All the samples were frozen in liquid nitrogen and save under −80°C before RNA extraction.

Subcellular Localization

The coding region of *GhSOC1-1* without stop codon was cloned from *G. hirsutum* TM-1 using gene-specific primers listed in **Supplementary Table 3** and fused with GFP driven by 35S promoter. The resulting vector was transformed into *Agrobacterium* by electroporation. The *Agrobacterium* harboring vector *35S:GhSOC1-1-GFP* were infiltrated into tobacco leaves. And the GFP signal was observed under confocal microscope Olympus FV1000.

Construction of Overexpression Lines

The full length CDS were cloned and constructed to vector pGreen0229 with 35S promoter. The resulting construct *35S:GhSOC1-1* was subsequently transformed into wild type *Arabidopsis* plants *via Agrobacterium* mediated transformation using floral dip method. The transgenic plants were selected with basta for positive lines which were verified by PCR examination on genome DNA with primers 35Spro and pgp2 (**Supplementary Table 3**). The expression of transgene was examined in the leaves of positive lines. The flowering phenotype was observed among at least 25 individuals of each line.

Virus-Induced Gene Silencing

For the VIGS assay, specific primers were designed to clone the 404 bp fragment of *GhSOC1-1* including 340 bp C-termina and 64 bp 3′ UTR into the *PTRV2* vector. Then, *Agrobacterium* carrying the plasmids of *PTRV2:GhSOC1-1* and *PTRV1*, respectively, were co-infiltrated into cotyledons of NDM8 after sowing for 7 days. *TRV:CLA1* and *TRV:00* were utilized as positive and negative control, respectively. The leaves were collected from each *GhSOC1-1* silencing lines to detect silence efficiency when albinism was obvious on the positive control. The flowering time of silencing lines were record when the first square appeared.

qPCR Examination

The collected samples were grinded for total RNA extracted with RNA prep Pure Plant Kit (TIANGEN), followed by synthesis of cDNA using PrimeScript^{TM} 1st strand cDNA Synthesis kit (Solarbio). Gene transcription was detected with AugeGreen qPCR Master Mix (US Everbright) on an ABI Q5 machine. *GhHIS* and *AtTUB2* were set as the internal control. Gene specific primers for qPCR were listed in **Supplementary Table 3**. The relative expression level was calculated using the $2^{-\Delta CT}$ formular (Livak and Schmittgen, 2001). There biological repeats were applied on each sample and three technical repeats were performed on each reaction.

RESULTS

Structure Analysis of *SOC1*-Like Genes in *Gossypium hirsutum*

The protein sequence of AtSOC1 (AT2G45660) was obtained from TAIR[7] and blast against the cotton genomes on Cotton-FGD.[8] After CD-search for the conserved domains, six *SOC1*-like genes were confirmed in *G. hirsutum* considering the most recent genome of NDM8 and other five genomes of TM-1, which were named according to their chromosome location (**Supplementary Table 1**). The sequence similarity of homologs located in the A and D subgenomes was up to 95% (**Supplementary Table 5**), suggesting their redundant function between A and D subgenome. The length of GhSOC1 proteins varied little with only 23 amino acid difference, resulting in similar molecular weight (23.200–25.713 kDa). The isoelectric point of GhSOC1-3D was lower than the other GhSOC1s that showed similar pI within 9.01–9.33, which was related to the sequence variation (**Supplementary Table 4**). *GhSOC1-2A* and *GhSOC1-2D* were highly conserved with the published GhSOC1 gene (Zhang et al., 2016).

Gene structure analysis of *SOC1* homologs showed that they were composed of seven exons and six introns. The length of Exon 1, 4, 5 and 6 were highly conserved and varied within Exon 7. The variations of intron length contributed to the gene length (**Figure 1A**). Noticeably, *GhSOC1-1* contained a large Intron 1, which distinguished from *AtSOC1* and other *GhSOC1s*. Then the MADS and K-box domains were compared in detail. GhSOC1s contained highly conserved MADS-box domain and

[7]https://www.arabidopsis.org/

[8]https://cottonfgd.org/

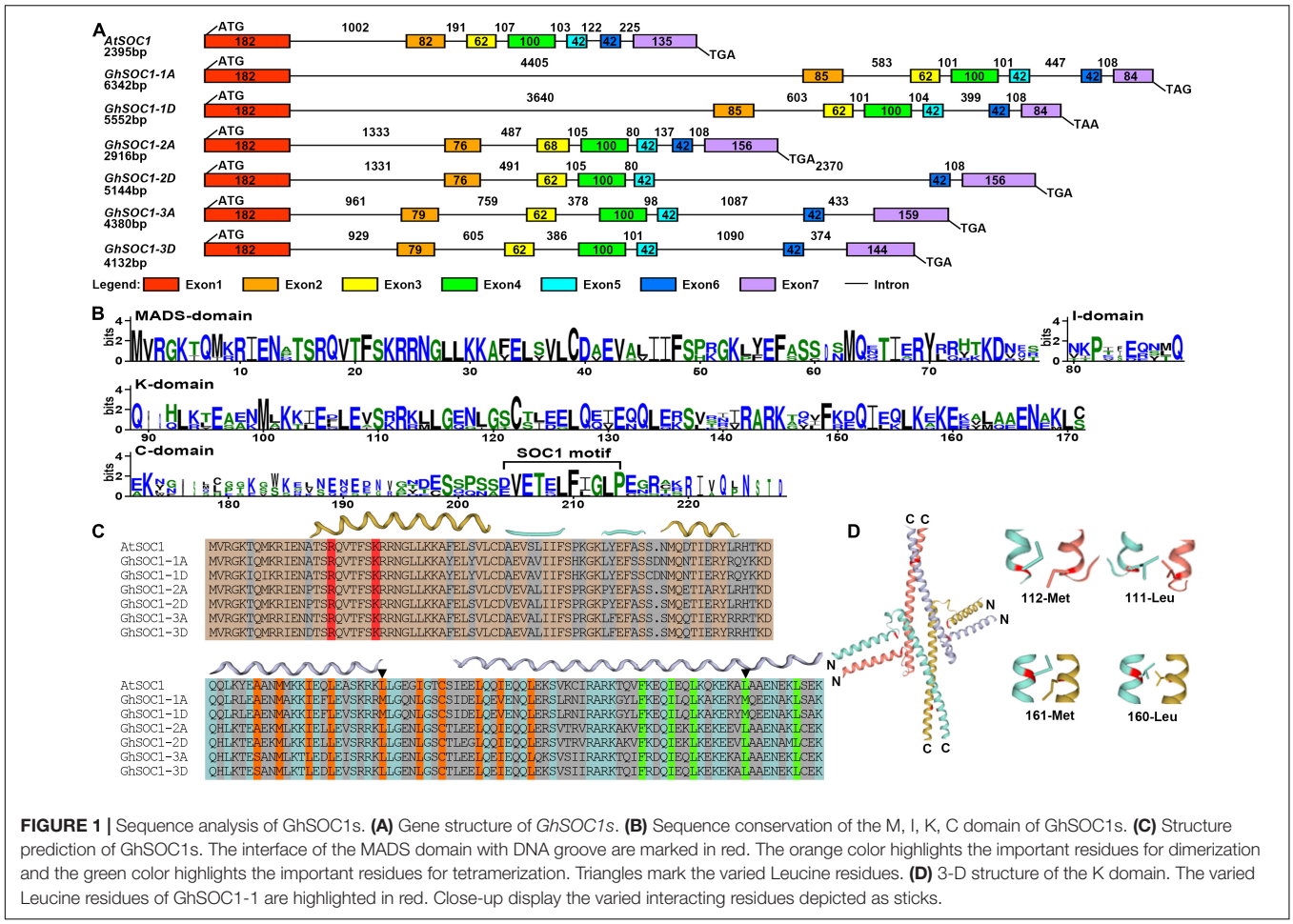


FIGURE 1 | Sequence analysis of GhSOC1s. **(A)** Gene structure of *GhSOC1s*. **(B)** Sequence conservation of the M, I, K, C domain of GhSOC1s. **(C)** Structure prediction of GhSOC1s. The interface of the MADS domain with DNA groove are marked in red. The orange color highlights the important residues for dimerization and the green color highlights the important residues for tetramerization. Triangles mark the varied Leucine residues. **(D)** 3-D structure of the K domain. The varied Leucine residues of GhSOC1-1 are highlighted in red. Close-up display the varied interacting residues depicted as sticks.

less conserved K-box domain, and diverse C terminal including a typical SOC1-motif (**Figure 1B**).

Oligomerization is common for MADS transcription factors to recognize different targets. The K domain is critical for dimerization and tetramerization capabilities. The K domain contains two helices that helix 1 stabilized the tetramer facilitated by helix 2 (Puranik et al., 2014). We then predicted the structure of GhSOC1s and found that the MADS domain interface with DNA groove were residue Arginine and Lysine that were conserved among GhSOC1s (**Figure 1C**). It has been reported that the Leucine residues in the K domain are essential for co-operative DNA binding of MADS proteins (Rumpler et al., 2018). Two variations of the Leucine were identified in GhSOC1-1 (**Figure 1C**). The variation 112-met was at the end of Helix 1 that might affect dimerization. The variation 161-met comprised the tetramerization interface (**Figure 1D**). The variations of Leucine indicated the different binding sites or binding affinity of GhSOC1-1.

Phylogenetic Analysis of Cotton *SOC1*-Like Genes

In 2020, it was reported that *G. hirsutum* might be originated from the genome polyploidization of *G. raimondi* and A_0 (Huang et al., 2020). To study the evolution of *SOC1*-like genes, we further identified SOC1 homologies in *G. raimondii*, *G. arboreum* and the other tetraploid cotton *G. barbadense*. Then other SOC1-like genes were obtained from dicots (*Arabidopsis thaliana*, *Actinidia chinensis*, *Lactuca sativa*, *Vitis vinifera*, *Fragaria vesca*, *Fragaria* × *ananassa*, *Medicago truncatula*, *Glycine max*, *Orchidaceae*, *Theobroma cacao*, *Brassica rapa*, *Nelumbo nucifera*) and monocots (*Oryza sativa*, *Sorghum bicolor*, *Zea mays*, *Musa nana Lour.*, *Triticum aestivum*, *Saccharum spontaneum*) (**Supplementary Table 2**). The sequences of the homologies were analyzed in detail using neighbor-joint method to build an unrooted phylogenetic tree (**Figure 2**).

Phylogenetic analysis showed that cotton *SOC1*-like genes were more conserved within the A genomes, but evolved differently between D genomes. The phylogenetic tree was mainly clustered into two groups. Cotton *SOC1-2* and *SOC1-3* genes were clustered with other SOC1-like genes in dicots. However, cotton *SOC1-1* genes together with monocot *SOC1* genes formed a monophyletic group (**Figure 2**). Lotus is an ancient dicot, whose *SOC1*-like genes were also clustered with the monocot *SOC1* genes. Cacao is considered as an evolutionary close specie with Gossypium, whose genome contained two *SOC1*-like genes that one was conserved with monocot *SOC1* genes and the other evolved with dicot *SOC1* genes. These findings suggested that

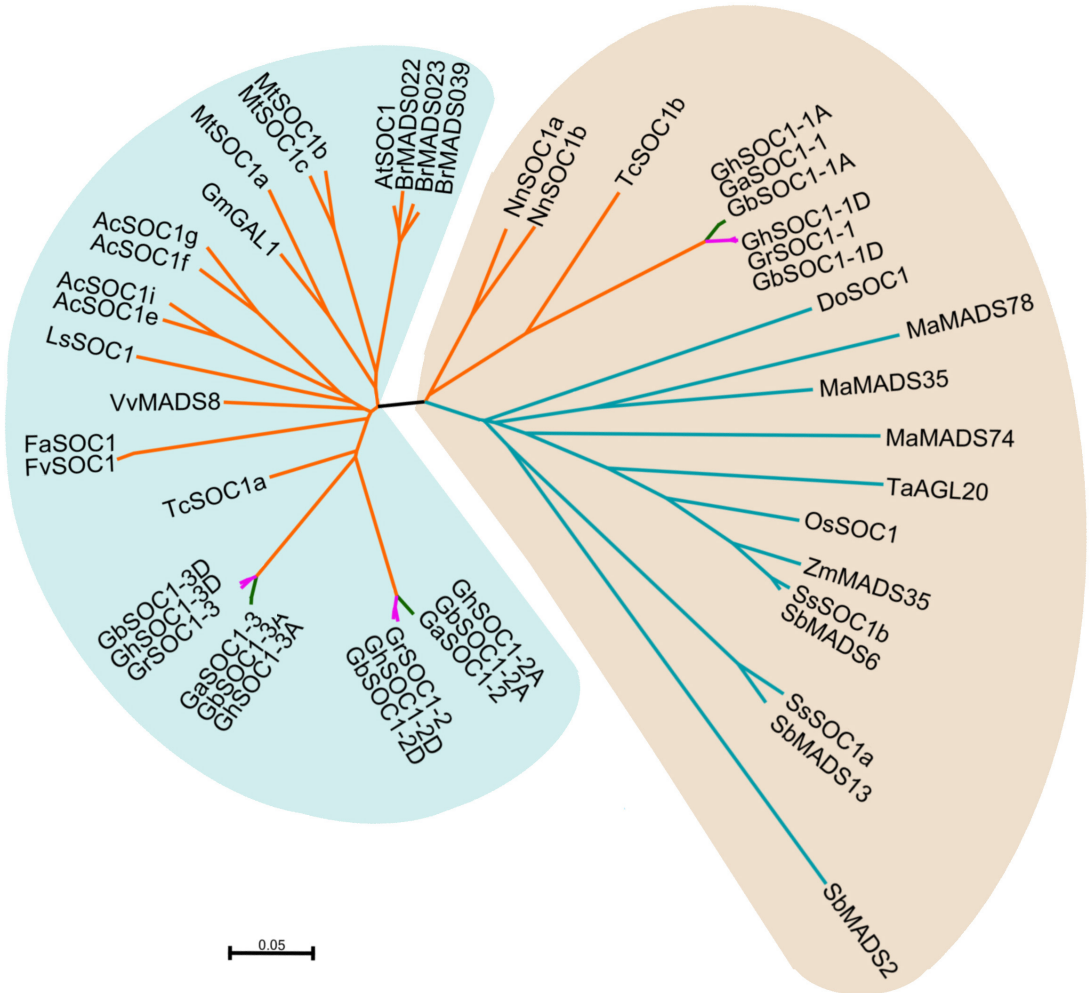


FIGURE 2 | Phylogenetic analysis of *SOC1*-like genes in cotton and other plants. SsSOC1a and SsSOC1b in *Saccharum spontaneum* was download from the article (Fatima et al., 2020). Other SOC1 sequences were downloaded from NCBI and their accession numbers are given in **Supplementary Table 2**.

Gossypium genomes retained the conserved ancient *SOC1* gene and evolved other *SOC1*-like genes with dicots.

Prediction of Regulatory Elements on Genomic Regions of *GhSOC1s*

Cis-elements are the binding sites for transcription regulators. We predicted the cis-elements on the 2 kb upstream sequences of *GhSOC1s*. The promoters of *GhSOC1s* contained plenty of light responsive elements, including 3-AF1 binding site, ABRE, AE-box, AT1-motif, ATCT-motif, Box 4, chs-CMA1a, GATA-motif, G-box, GT1-motif, I-box, MRE and CT-motif. The *GhSOC1-3D* promoter harbored unique elements relating to photoperiod (circadian). A low temperature responsive element LTR was found on promoters of *GhSOC1s*. Salicylic acid and auxin responsive elements were predicted on *GhSOC1-1/2/3* promoter region. *GhSOC1-1* and *GhSOC1-2* were predicted to response to abscisic acid. *GhSOC1-2* and *GhSOC1-3* might be regulated by gibberellin. And jasmonate responsive elements only existed on *GhSOC1-1* promoters. *SOC1*-like genes are important regulators in plant growth and development which could be disturbed and interrupted by biotic and abiotic stresses. Although little is known for *SOC1*-like genes in response to stress, TC-rich repeats that are involved in defense and stress response were identified on *GhSOC1-1* and *GhSOC1-2* promoter.

Since the *GhSOC1-1* genome comprised a large Intron 1, we further compared the cis-elements located in this region. The results showed that light responsive elements were all predicted in Intron 1 of *GhSOC1s*, but *GhSOC1-1* possessed more abundant regulatory elements in Intron 1 (**Figure 3**). This region of *GhSOC1-2* included JA and drought regulatory elements, whereas ABA, SA, GA and low-temperature responsive elements were also found in Intron 1 of *GhSOC1-1*. The large number of cis-elements indicated regulatory importance of Intron 1 in *GhSOC1-1* genome.

The diversity of *cis*-elements on genome region of *GhSOC1s* suggested differences in transcription and regulatory mechanism. The existence of abundant light related elements indicated

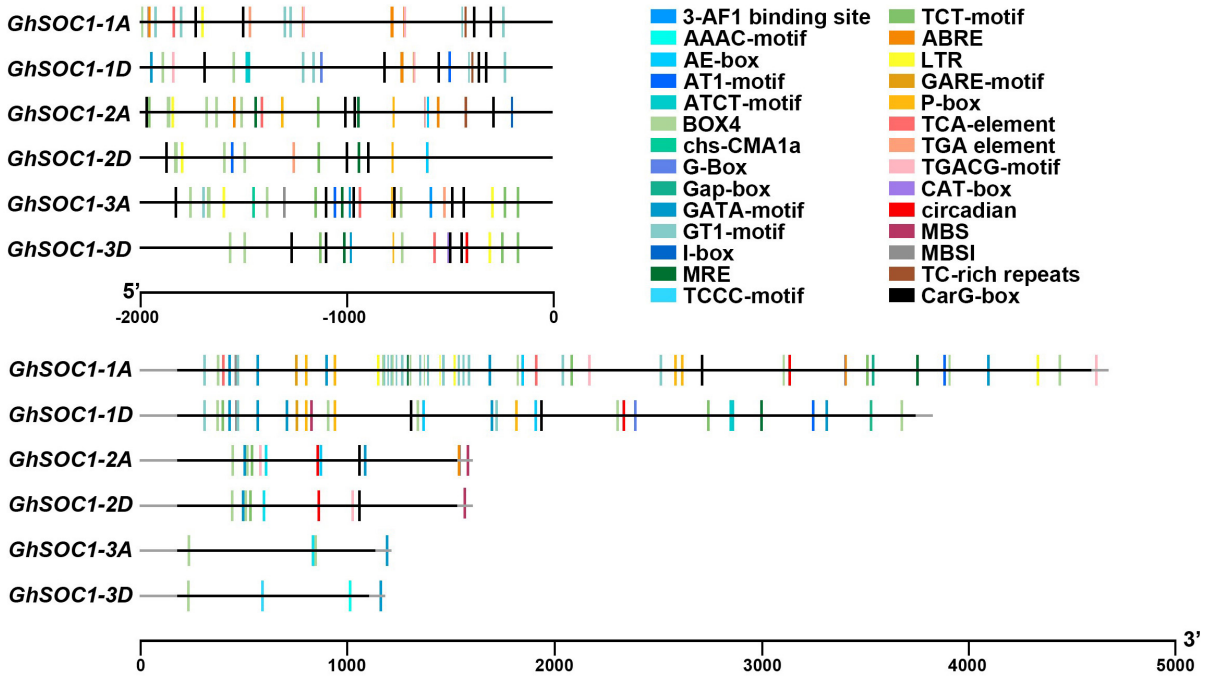


FIGURE 3 | Prediction of *cis*-acting elements in *GhSOC1s* genome. The coding region and non-coding region are presented in gray and black lines.

a regulation of *GhSOC1s* by light effectors, which might contribute to the loss of light sensitivity of upland cotton during domestication.

Expression Analysis of *GhSOC1s*

Expression of *GhSOC1s* were examined in roots, stems, leaves, flower and fiber. *GhSOC1s* expressed ubiquitously in different organs (**Figure 4A**). *GhSOC1-1* transcripts were most abundant in all the examined tissues (**Figures 4A,B**). *GhSOC1-3* was transcribed less abundantly with relatively high expression in the vegetative organs. The expression of *GhSOC1-1* and *GhSOC1-2* was higher in flowers. Therefore, *GhSOC1s* transcriptions were compared between different parts of flowers according to the transcriptome data (**Figure 4B**). Both *GhSOC1-1* and *GhSOC1-2* were expressed higher in the calyx, and *GhSOC1-3* demonstrated an expression preference in torus. *GhSOC1-1A* also expressed in the petals, stamens and pistils. Then the expression patterns were detected in developing seedlings. The *GhSOC1s* expression increased gradually during seedling development (**Figure 4C**). Especially, their expression was upregulated sharply at the fourth true leaf stage, indicating the occurrence of floral transition.

Further, the coding region of *GhSOC1-1* was cloned from TM-1 because it was clustered in different branch from other *GhSOC1s* in the phylogenetic tree and its transcript was most abundant. The amplification displayed the same sequence as *GhSOC1-1* in genome of NDM8 (Ma et al., 2021). The coding region was subsequently ligated with GFP sequence to construct GhSOC1-1-GFP fusion protein. The fluorescence was observed in the nucleus of the tobacco epidermal cells (**Figure 4D**), which distinguished from AtSOC1 that only localized in the nuclear with existence of AtAGL24 (Lee et al., 2008). The nuclear localization of GhSOC1-1 supported its role as a transcription factor.

Overexpression of *GhSOC1-1* Promotes Flowering

In order to study the function of *GhSOC1-1*, *35S:GhSOC1-1* vector was constructed, and transformed into *Arabidopsis* wild-type plants. The positive lines were confirmed by amplification of the exogenous gene from genome DNA (**Figure 5A**). Finally, twenty-nine individual transgenic lines were obtained that displayed consistent early flowering under long day condition (**Figure 5B**). Four transgenic lines were randomly selected for *GhSOC1-1* expression detection. The expression levels related with the number of rosette leaves (**Figure 5C**), suggesting that *GhSOC1-1* functioned in a dosage-dependent way to promote flowering. The function of *GhSOC1-1* in flowering time control was consistent with its temporal expression pattern.

SOC1, *FT*, and *LFY* are floral integrator in the network of flowering time control. Therefore, the expression of *FT* and *LFY* were detected in overexpression lines to investigate their regulatory relationship. The results showed that overexpression of *GhSOC1-1* did not affect *FT* and AP1 expression, while *LFY* expression was greatly upregulated (**Figure 5D**).

Silencing *GhSOC1-1* Delays Flowering

To further confirm the function of *GhSOC1-1* in cotton, the Virus-induced gene silencing (VIGS) was employed using *TRV* containing a fragment of the C terminal of *GhSOC1-1* and silencing of *CLA1* provided a visible reporter of silencing

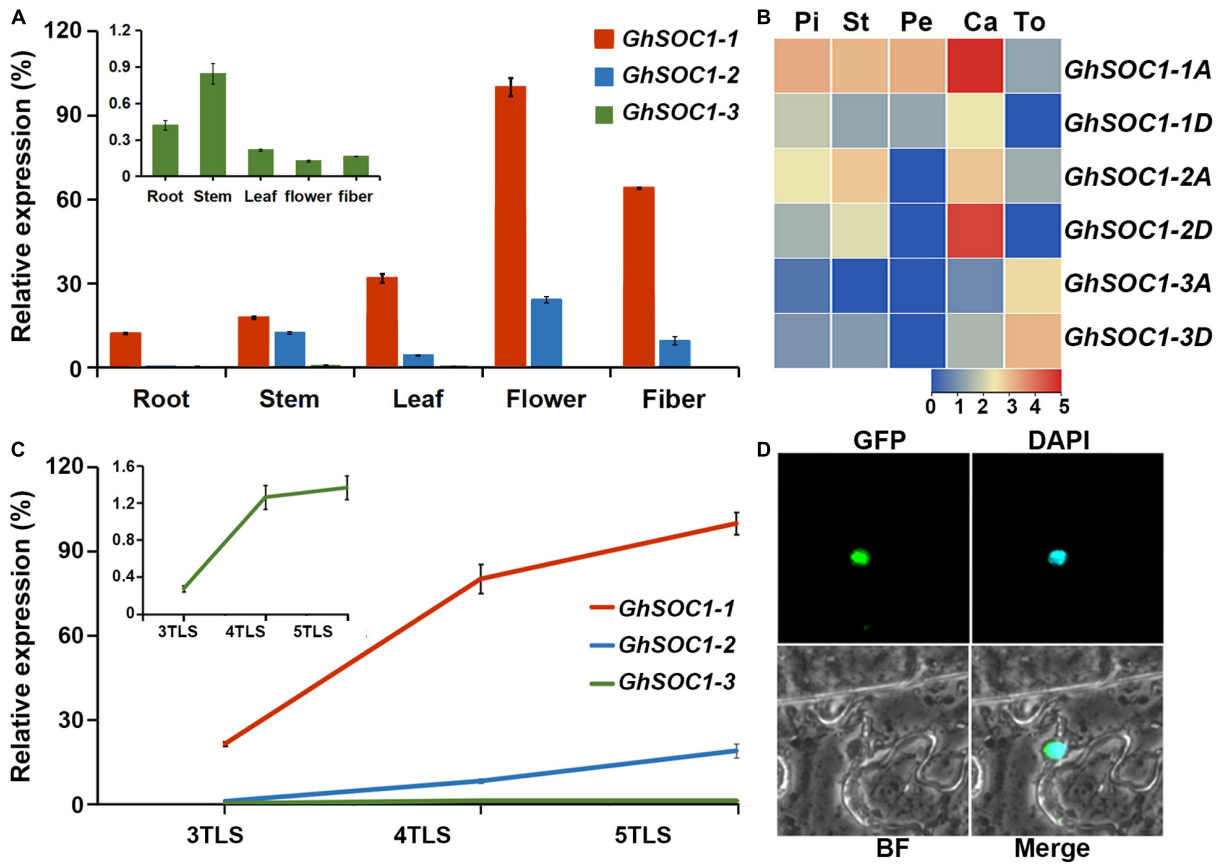


FIGURE 4 | Expression analysis of *GhSOC1s*. **(A)** Tissue specific expression of *GhSOC1s* in *G. hirsutum*. **(B)** Expression pattern of *GhSOC1s* in a flower. Pi, pistil; St, stamen; Pe, petal; Ca; calyx; To, torus. **(C)** Expression changes of *GhSOC1s* during seedling development. 3TLS, 4TLS, and 5TLS demonstrates the third, fourth and fifth true leaf stages, respectively. The expression is normalized to *GhHIS* and the maximum expression is set as 100%. Error bars represent the standard deviations of three biological replicates. **(D)** Subcellular localization of GhSOC1-1 in tobacco leaves. GFP, GFP fluorescence; DAPI, fluorescence of 40,6-diamino-2-phenylindole; BF, Bright field; Merge, merge of GFP, DAPI, and BF images.

effects. The silence efficiency was detected when the *TRV:CLA1* plants demonstrated obvious leaf photobleaching (**Figure 6A**). *GhSOC1-1* expression was only a quarter of that in the control plants infected with empty *TRV* vectors, meanwhile the expression of other *GhSOC1* genes also decreased by nearly half (**Figure 6C**), resulting from the sequence similarity. The first square appeared on the eighth branch of the control plants, while the square grew out on the tenth branch of the *GhSOC1*-silenced plants (**Figures 6B,D,E**). The flowering time was delayed significantly by downregulation of *GhSOC1s*, suggesting their consistent roles in promoting floral transition. Furthermore, the expression changes of *GhFT*, *GhAP1*, and *GhLFY* were consistent with the results of Arabidopsis that only *GhLFY* displayed significant decrease (**Figure 6F**).

GhSOC1s Response to Light and Ambient Temperature

Considering the existence of light and temperature responsive *cis*-elements in the genome region of *GhSOC1s*, we examined their transcriptional level under different environmental conditions. Firstly, *GhSOC1-3* expression was lowest under long-day condition and *GhSOC1-1* transcription was higher than *GhSOC1-2* (**Figure 4A**). When the plants were shifted to short-day for 48 h, their expression remained unchanged (**Figure 7A**). Differently, *GhSOC1-1* and *GhSOC1-2* responded to temperature change quickly (48 h), and their response to high temperature were violent (**Figure 7B**). But *GhSOC1-3* didn't response to temperature change in a short time. These results suggested that *GhSOC1-1* and *GhSOC1-2* might involve in the ambient temperature pathway to regulate flowering time.

Secondly, the expression analysis suggested the occurrence of floral transition of TM-1 at 4TLS (**Figure 4C**), and thus *GhSOC1s* expression changes before and after floral transition in response to photoperiod and ambient temperature were compared at 3TLS and 5TLS. The results showed that *GhSOC1-1* expression elevated dramatically under short-day during floral transition, while expression elevation of *GhSOC1-2* and *GhSOC1-3* were more violent under long-day compared with short-day, suggesting that *GhSOC1-2* and *GhSOC1-3* were more influential to floral transition in response to long-day condition (**Figure 7C**). The differences of expression elevation during

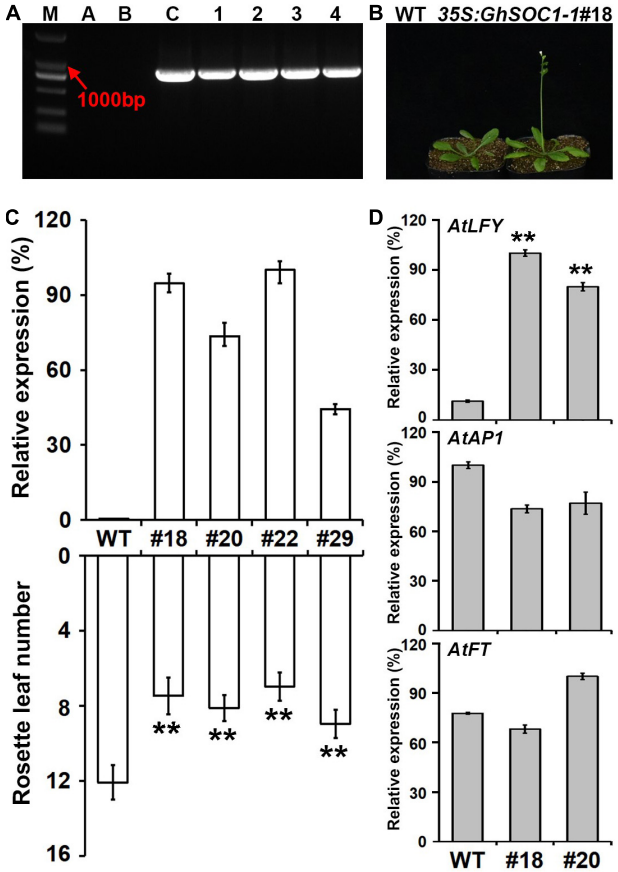


FIGURE 5 | Overexpression of *GhSOC1-1* promotes flowering of *Arabidopsis*. **(A)** Genotype verification of positive plants in T1 generation. M, DNA molecular weight marker DL 2000; A-C, water control, WT control and positive control; 1-4, PCR detection of positive plants. **(B)** Phenotype of T3 generation line. **(C)** Relative expression of *GhSOC1-1* and statistics of flowering time in transgenic lines and wild-type *Arabidopsis*. Asterisks indicate statistically significant differences according to Student's *t*-test (**$P < 0.01$). **(D)** Expression of AtLFY, AtAP1, and AtFT in transgenic and wild-type Arabidopsis. The expression is normalized to *AtTUB2* and the maximum expression is set as 100%. Error bars represent the standard deviations of three biological replicates.

floral transition revealed distinct participation of *GhSOC1s* in response to photoperiod.

Thirdly, the responses of *GhSOC1s* to ambient temperature were different from the quick response to temperature change. *GhSOC1s* expression under high temperature was the lowest compared to normal temperature (28°C) and only *GhSOC1-1* expression significantly increased during floral transition under 35°C (**Figure 7D**), suggesting that *GhSOC1-1* rather than *GhSOC1-2/3* acts as an effector of high ambient temperature to accelerate flowering. Under 18 and 22°C, *GhSOC1s* expression was promoted at the vegetative growth stage and *GhSOC1-3* increased significantly during floral transition. Transcription of *GhSOC1-2* was dramatically induced under 18°C, suggesting a unique role of *GhSOC1-2* in response to cold. It was likely that *GhSOC1s* played cooperatively under low ambient temperature to promote flowering and ensure the success of reproductive growth.

DISCUSSION

Gossypium origins from the tropical and subtropical area that is adapted to short-day photoperiod and the megathermal climate. The cultivated species has changed after domestication. The widely planted upland cotton losses the seed dormancy and sensitivity to short-day photoperiod and produces more fibers with increased length and whiteness (Chen et al., 2020). The application of GWAS has revealed the molecular mechanism of domestication on some traits and identified some locus contributing to maturity (Ma et al., 2018; Li et al., 2021a). However, our knowledge on cotton flowering time control is limited, restricting the co-improvement of cotton on maturity, yield, quality and resistance.

GhSOC1s Diverged After γ Whole-Genome Duplication

Previous studies have identified several *SOC1* homologs in tetraploid upland cotton (Nardeli et al., 2018), one of which was cloned as *GhSOC1* (*GhSOC1-2*). In this study, six *SOC1* homologous were identified in upland cotton genome with the conserved domains (MADS-box and K-box) and SOC1-motif (**Figure 1B**). But they formed two clades in the phylogenetic tree that *GhSOC1-1* sequences were more similar to monocot *SOC1* sequences and clustered in the different group from the dicot *SOC1* subclade including *GhSOC1-2* and *GhSOC1-3* (**Figure 2**). Consistent with the phylogenetic tree, the gene structure of *GhSOC1-1* characterized by a large first intron is similar as *OsSOC1* (Tadege et al., 2003).

Whole-genome duplication (WGDs) is significant for transformative evolution. The gamma (γ) WGD occurred after the derivation of monocots from the core dicots (Vekemans et al., 2012). The *Vitis* species only experienced this ancient whole-genome triplication, during which *SOC1* genes triplicated (Vekemans et al., 2012; Zhang et al., 2020). Cocoa, the lineage close specie of *Gossypium*, also experienced the γ WGD. But, the early diverging dicot lineage of lotus (*Nelumbo nucifera*) did not experience this whole-genome triplication (Zhang et al., 2020). Therefore, cotton *SOC1-1* genes that were close to *NnSOC1* and monocot *SOC1* genes should be the ancient loci retained in the Gossypium genomes, and triplicated during the γ WGD after which cotton *SOC1-2* and *SOC1-3* genes evolved as the dicot *SOC1* genes. The conserved gene structure further supported the above inference.

Conservation and Diversification of *GhSOC1s* Function

The early flowering phenotype caused by overexpression of *GhSOC1-1* (**Figures 5B,C**) or *GhSOC1-2* (Zhang et al., 2016) in *Arabidopsis* suggested functional conservation in flowering time control. And the consistent sharp increase of *GhSOC1s* during the occurrence of floral transition (**Figure 4C**) further suggested their conserved function in promoting flowering. Corporate silencing

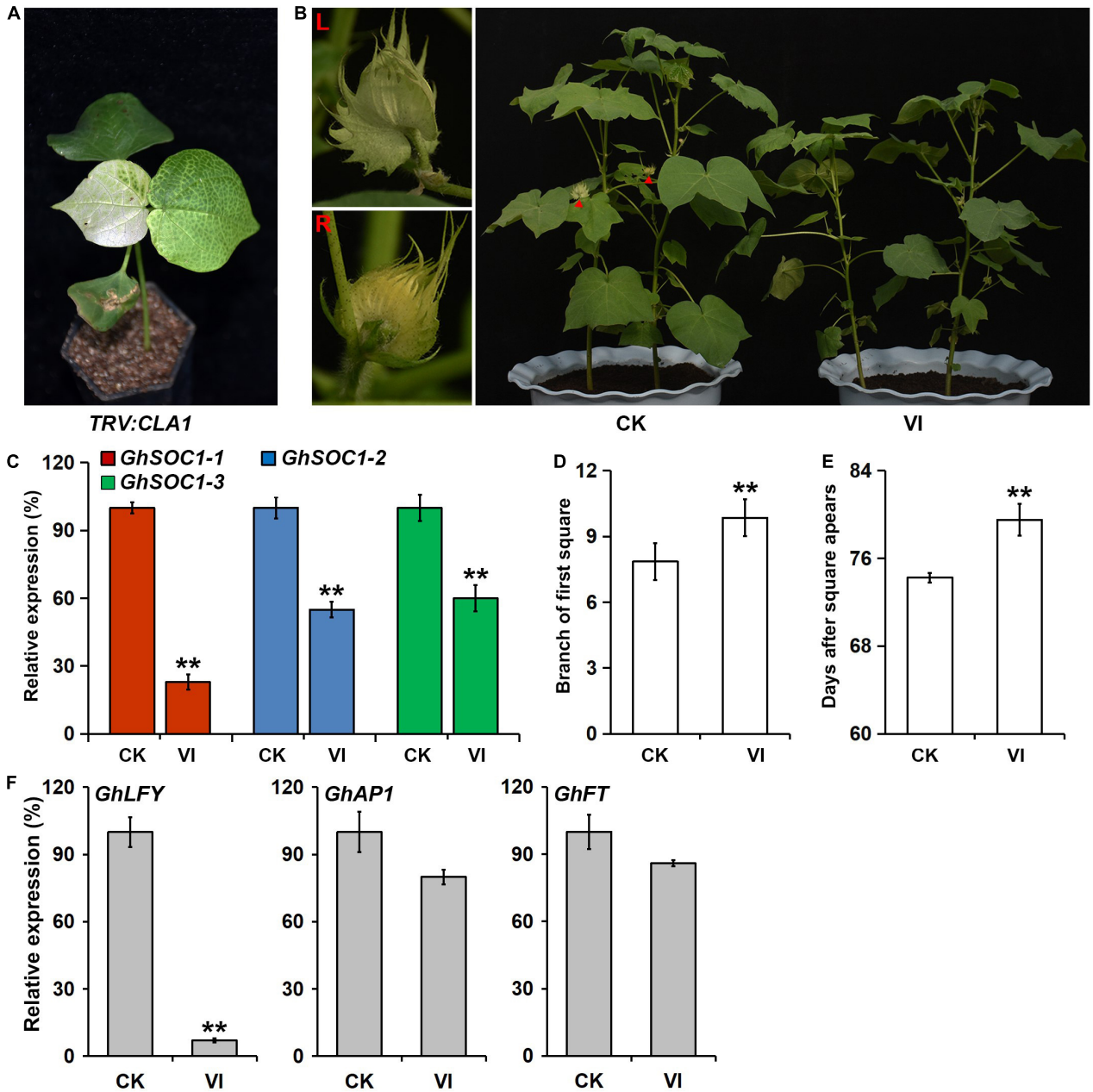


FIGURE 6 | VIGS of *GhSOC1s* in *G. hirsutum* delays flowering time. **(A)** Albino phenotype of *TRV:CLA1* as a control. **(B)** Delayed flowering phenotype of VIGS plants. Red triangles point the flower buds displayed in the left. CK, empty vector control; VI, silence plants. **(C)** Detection of the efficiency of gene silencing. The expression was normalized to *GhHIS* and the maximum expression is set as 100%. Error bars represent the standard deviations of three biological replicates. **(D,E)** Maturity traits of VIGS plants in branch of first square **(D)** and days from sowing until square appears **(E)**. Asterisks indicate statistically significant differences among 15 individuals according to Student's *t*-test (**$P < 0.01$). **(F)** Expression of *GhLFY*, *GhAP1,* and *GhFT* in Silencing and control plants. The expression is normalized to *GhHIS* and the maximum expression is set as 100%. Error bars represent the standard deviations of three biological replicates.

of *GhSOC1-1/2/3* delayed cotton flowering, confirming the functional overlap of *GhSOC1s* as flowering promoters. Another commonality was observed in their ubiquitous expression in vegetative organs (**Figure 4A**), which is consistent with other *SOC1* homologs as flowering regulators (Lee et al., 2000; Wei et al., 2016; Liu et al., 2020). Moreover, SOC1 act as a floral integrator to function in the downstream of the flowering time regulation network to mainly activate LFY expression (Lee and Lee, 2010). The expression changes of *LFY* in both *GhSOC1-1* overexpression and silencing plants revealed that *GhSOC1-1* promote flowering *via GhLFY*.

However, the diversity of *cis*-elements in the promoter of different *GhSOC1s* suggest discrepancy of their transcription regulation (**Figure 3**), which was supported by their expression

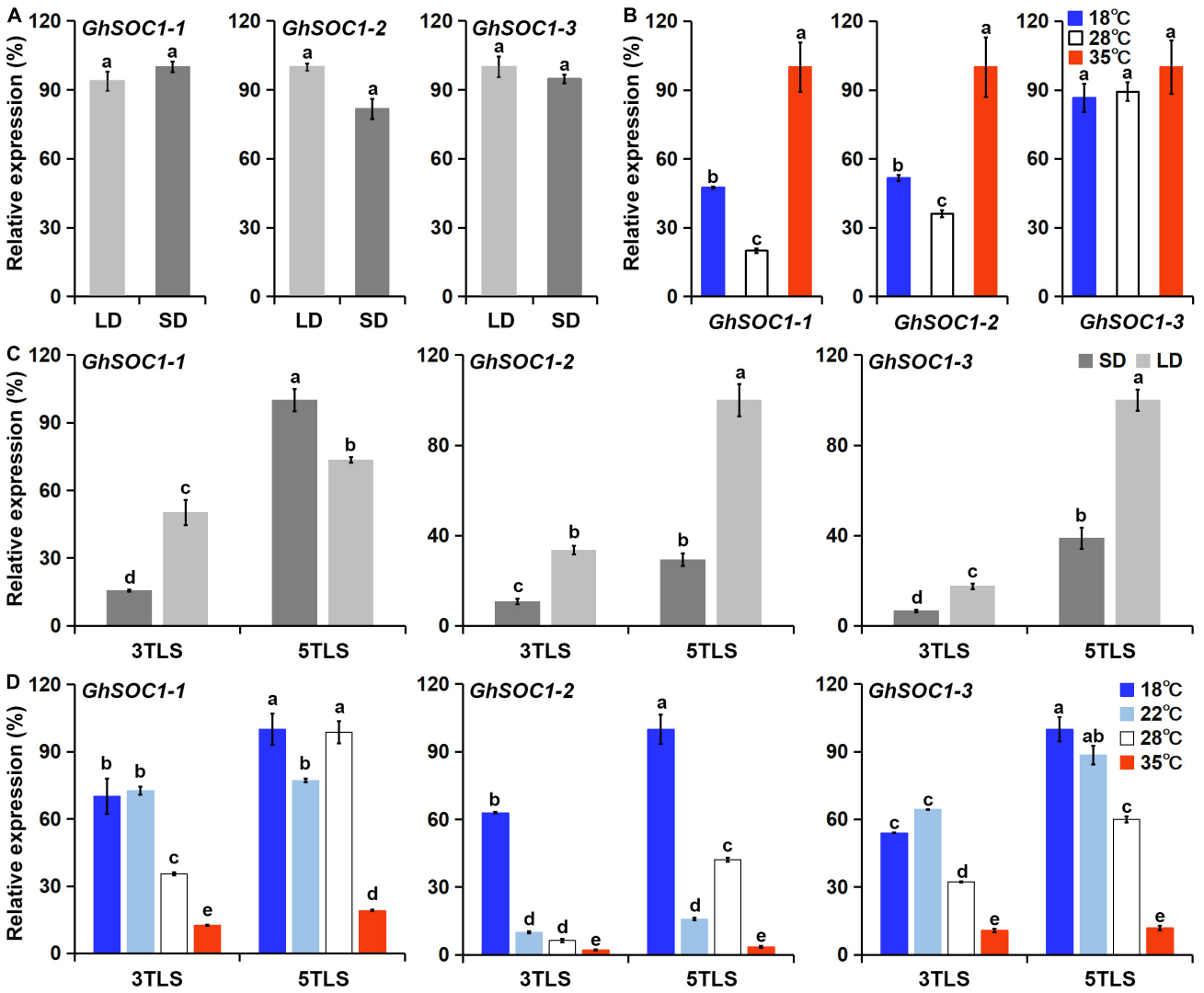


FIGURE 7 | Responses of *GhSOC1s* to light and temperature. **(A)** Quick response of *GhSOC1s* to day length shift from long-day to short-day for 48 h under 28°C. **(B)** Quick response of *GhSOC1s* to temperature shift for 48 h examined in seedlings grown under 28°C and long-day condition. **(C)** *GhSOC1s* expression is examined under continues long-day and short-day conditions under 28°C at 3TLS and 5TLS. **(D)** *GhSOC1s* expression is examined under different temperature conditions under long-day at 3TLS and 5TLS. The expression is normalized to *GhHIS*. Error bars represent the standard deviations of three biological replicates. Different letters suggest significant differences calculated by Student's *t*-test ($P < 0.01$).

differences (**Figure 4A**). First, their transcription abundance varied greatly. The actively transcribed *GhSOC1-1* may play a dominant role, while minimal expression of *GhSOC1-3* was detected prevalently in roots and stems indicating a functional divergence. Although the coding region determines gene function, the introns are proved to contain *cis*-acting elements. For example, the 3.5 kb first intron of *AtFLC* are critical for the epigenetic repression (Sheldon et al., 2002). The large intron might contribute to the abundant transcription of *GhSOC1-1* (**Figure 1A**). The failure in promoting flowering of constitutive expression of *GhSOC1-2* in cotton is evidence for functional divergency in flowering time control (Zhang et al., 2016). Moreover, ABF3 and ABF4 bind to SOC1 promotor to mediate drought-accelerated flowering (Hwang et al., 2019). The ABRE elements were predicted in *GhSOC1-1* and *GhSOC1-2* regulatory region, but drought responsive elements only existed in *GhSOC1-1* genome (**Figure 3**). As functions of ABFs diverge and overlap (Finkelstein et al., 2005), GhSOC1-1 with more ABRE elements was more likely to respond to drought to promote flowering.

GhSOC1s expressed differently in floral organs. *GhSOC1-3* showed higher expression in vegetative organs, but *GhSOC1-1* and *GhSOC1-2* were most abundantly expressed in flower. Detailed expression of *GhSOC1s* in flowers demonstrated that they were accordantly expressed highest in calycale, and *GhSOC1-1* also accumulated in the inner whorls, indicating their participation in flower development. It was supported that *35S:GhSOC1-2* causes floral defects in cotton (Zhang et al., 2016).

In contrary, the cotton plants produced normal flowers when *GhSOC1s* were silenced (data not shown), which is the same as soc1-2 mutant of *Arabidopsis*. But triple mutants of *soc1 svp agl24* display sever floral defects (Liu et al., 2009). The *Gerbera hybrida SOC1*-like gene is only expressed in flowers and overexpression caused partial defects of petal on color and shape without changes in flowering time (Ruokolainen et al., 2011). Therefore, *GhSOC1s* might be involved in the development of different floral organs with other MADS cooperators. Further, the function in floral development needs to be induces. As *FBP21/22* and *AtSOC1* affects petal color and development only under high temperature (Wang et al., 2019).

Additionally, quick response of *GhSOC1-1* and *GhSOC1-2* to high and low temperature indicated involvement in stress resistance (**Figure 7B**), which is supported by the existence of ABA responsive and defense related cis-elements in their genome rather than *GhSOC1-3* (**Figure 3**).

GhSOC1s Respond Differently to Environmental Cues to Promote Flowering

In the model plant, *SOC1* integrates the environmental and endogenous signals to initiate floral transition (Lee and Lee, 2010). And *SOC1* homologs diverged in role of flowering time control among species. *TgSOC1-like1* and *TgSOC1-like2* functions opposite in flowering time control (Leeggangers et al., 2017). Four *SOC1*-like genes (*AcSOC1e*, *AcSOC1f*, *AcSOC1i*, *AcSOC1g*) from Actinidia sinensis promote flowering of Arabidopsis in varying degrees, but respond differently to winter chilling to regulate woody perennials (Voogd et al., 2015). The Medicago truncatula genome possesses three *SOC1* genes (*MtSOC1a—MtSOC1c*) that are expressed differently in response to day-length and vernalization to promote flowering (Fudge et al., 2018). Similarly, *GhSOC1s* respond differently to photoperiod although their genomes all possess a large number of light responsive elements (**Figures 3**, **7C**). *GhSOC1-1* was greatly induced by short-day condition, while *GhSOC1-2* and *GhSOC1-3* respond to long-day. These suggested that *GhSOC1-1* evolved conserved as discussed above, while *GhSOC1-2* and *GhSOC1-3* contribute to the insensitivity of day-length during domestication.

Ambient temperature is another important environmental cue that affects flowering time. Cotton can grow normally under the temperature above 22°C. Low ambient temperature induced expression of *GhSOC1s* in the vegetative stage, and *GhSOC1-3* expression further elevated during floral transition (**Figure 7D**). But only *GhSOC1-1* respond to the high ambient temperature to activate floral transition (**Figure 7D**). The response of *GhSOC1-1* to high ambient temperature together with the induction by short-day is reasonable for the dominant role in flowering time control as cotton origins in the subtropical area. When the temperature is not suitable for cotton growth, *GhSOC1-2* was dramatically induced. And overexpression of *GhSOC1-2* causes significant increase of *FT* and *LFY* expression (Zhang et al., 2016). We speculated that *GhSOC1s* act cooperatively to promote flowering *via GhFT* and *GhLFY* (**Figure 8**), which would be essential for seed generation under different growth conditions to ensure reproductive success.

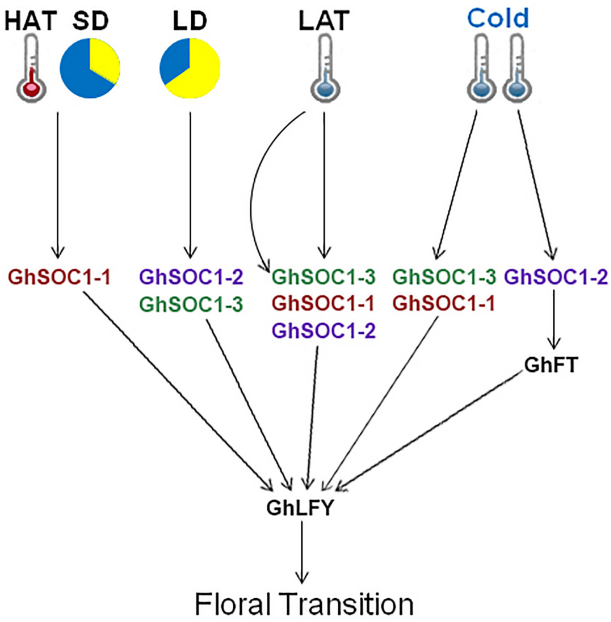


FIGURE 8 | Proposed regulatory model of *GhSOC1s*. HAT, high ambient temperature; LAT, low ambient temperature; LD, long-day condition; SD, short-day condition.

Multimerization is common for MADS proteins contributing to a large protein-protein interaction network that fine tune the reproductive growth (Theissen et al., 2016). The formation of SOC1-AGL24 heterodimer is compulsory for entrance of SOC1 to the nucleus where SOC1-AGl24 binds to the target gene (Lee et al., 2008). Similarly, expression of *FaSOC1* (Lei et al., 2013) or *GmGAL1* (Zhong et al., 2012) (*SOC1* homolog gene) alone, the resulting proteins are localized in the cytoplasm. Differently, GhSOC1-1 proteins were localized in the nucleus without any assistance to play its function (**Figure 4D**), which is consistent with LsSOC1 (Chen et al., 2018). Structure prediction found two variations of the essential Leucine residues in the K domain of GhSOC1-1 (**Figure 1D**). Single mutation at L115 and L164 is able to destroy the coiled-coil structure of the K domain, and one substitution mutation L164P or L115P reduces the ability to form tetramer resulting in altered DNA-binding, which is likely to be overcame by high protein concentration (Rumpler et al., 2018). Thus the interactions of GhSOC1-1 probably distinct from GhSOC1-2 and GhSOC1-3, and proper function of GhSOC1-1 depends on the concentration that is consistent with its abundant transcripts. The structure effects on DNA-binding provide an explanation for differences of the downstream gene. It has reported that GhSOC1-2 directly binds to the promotor of *GhLFY* to activate its transcription (Li et al., 2013). Overexpression of *35S:GhSOC1-1* or *35S:GhSOC1-2* upregulated *LFY* expression (**Figure 5D**; Zhang et al., 2016). Meanwhile, *FT* expression is greatly enhanced by overexpression of *GhSOC1-2* (Zhang et al., 2016). However, *FT* transcription was not affected by *GhSOC1-1*

(**Figure 5D**), suggesting that *GhSOC1-1* is the downstream of *FT* as in Arabidopsis. Thus, our results revealed that *GhSOC1-1* acts downstream of *FT* to promote floral transition *via LFY*. Although *GhSOC1-2* promotes *FT* expression in Arabidopsis, it failed to bind to *FT* genome (Zhang et al., 2016), suggesting another regulation pathway for *GhSOC1-2* (**Figure 8**).

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

YY conceived and designed the experiment. LM performed the experiments. YY and LM analyzed the data and wrote the manuscript. Both authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant no. 31801410), the Provincial Natural Science Foundation of Hebei (grant no. C2020204079), the Supporting Project of Hebei Agricultural University (grant nos. ZD201601 and PT2018004), and the Top Talent Project of Hebei Province to Prof. Ma Zhiying (031601801). We are grateful to Prof. Ma Zhiying for his advise and helps on this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.882946/full#supplementary-material

REFERENCES

Amrouk, E. M., Mermigkas, G., and Townsend, T. (2021). *Recent Trends and Prospects in the World Cotton Market and Policy Developments*. Rome: Food and Agriculture Organization of the United Nations, 72.

Borner, R., Kampmann, G., Chandler, J., Gleiûner, R., Wisman, E., Apel, K., et al. (2008). A MADS domain gene involved in the transition to flowering in *Arabidopsis*. *Plant J.* 24, 591–599. doi: 10.1046/j.1365-313x.2000.00906.x

Chen, Z., Zhao, W., Ge, D., Han, Y., Ning, K., Luo, C., et al. (2018). LCM-seq reveals the crucial role of LsSOC1 in heat-promoted bolting of lettuce (*Lactuca sativa* L.). *Plant J.* 95, 516–528. doi: 10.1111/tpj.13968

Chen, Z. J., Sreedasyam, A., Ando, A., Song, Q., De Santiago, L. M., Hulse-Kemp, A. M., et al. (2020). Genomic diversifications of five Gossypium allopolyploid species and their impact on cotton improvement. *Nat. Genet.* 52, 525–533. doi: 10.1038/s41588-020-0614-5

Ding, L., Wang, Y., and Yu, H. (2013). Overexpression of DOSOC1, an ortholog of Arabidopsis SOC1, promotes flowering in the orchid Dendrobium Chao Parya Smile. *Plant Cell Physiol.* 54, 595–608. doi: 10.1093/pcp/pct026

Fatima, M., Zhang, X., Lin, J., Zhou, P., Zhou, D., and Ming, R. (2020). Expression profiling of MADS-box gene family revealed its role in vegetative development and stem ripening in S. spontaneum. *Sci. Rep.* 10:20536. doi: 10.1038/s41598-020-77375-6

Finkelstein, R., Gampala, S. S., Lynch, T. J., Thomas, T. L., and Rock, C. D. (2005). Redundant and distinct functions of the ABA response loci *ABA-INSENSITIVE*(*ABI*)*5* and *ABRE-BINDING FACTOR* (*ABF*)*3*. *Plant Mol. Biol.* 59, 253–267.

Fudge, J. B., Lee, R. H., Laurie, R. E., Mysore, K. S., Wen, J., Weller, J. L., et al. (2018). Medicago truncatula SOC1 genes are up-regulated by environmental cues that promote flowering. *Front. Plant Sci.* 9:496. doi: 10.3389/fpls.2018.00496

Hao, J. J., Yu, S. X., Ma, Q. X., Fan, S. L., and Song, M. Z. (2008). Inheritance of time of flowering in upland cotton under natural conditions. *Plant Breed.* 127, 383–390. doi: 10.1111/j.1439-0523.2007.01474.x

Huang, G., Wu, Z., Percy, R. G., Bai, M., Li, Y., Frelichowski, J. E., et al. (2020). Genome sequence of *Gossypium herbaceum* and genome updates of *Gossypium arboreum* and *Gossypium hirsutum* provide insights into cotton A-genome evolution. *Nat. Genet.* 52, 516–524. doi: 10.1038/s41588-020-0607-4

Hwang, K., Susila, H., Nasim, Z., Jung, J. Y., and Ahn, J. H. (2019). *Arabidopsis ABF3* and *ABF4* transcription factors act with the NF-YC complex to regulate *SOC1* expression and mediate drought-accelerated flowering. *Mol. Plant* 12, 489–505. doi: 10.1016/j.molp.2019.01.002

Immink, R. G., Pose, 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

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

Kimura, Y., Aoki, S., Ando, E., Kitatsuji, A., Watanabe, A., Ohnishi, M., et al. (2015). A flowering integrator, *SOC1*, affects stomatal opening in *Arabidopsis thaliana*. *Plant Cell Physiol.* 56, 640–649. doi: 10.1093/pcp/pcu214

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

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

Lee, J., Oh, M., Park, H., and Lee, I. (2008). SOC1 translocated to the nucleus by interaction with AGL24 directly regulates leafy. *Plant J.* 55, 832–843. doi: 10.1111/j.1365-313X.2008.03552.x

Leeggangers, H. A., Nijveen, H., Bigas, J. N., Hilhorst, H. W., and Immink, R. G. (2017). Molecular regulation of temperature-dependent floral induction in *Tulipa gesneriana*. *Plant Physiol.* 173, 1904–1919. doi: 10.1104/pp.16.01758

Lei, H. J., Yuan, H. Z., Liu, Y., Guo, X. W., Liao, X., Liu, L. L., et al. (2013). Identification and characterization of FaSOC1, a homolog of SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 from strawberry. *Gene* 531, 158–167. doi: 10.1016/j.gene.2013.09.036

Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., et al. (2008). A repressor complex governs the integration of flowering signals in Arabidopsis. *Dev. Cell* 15, 110–120. doi: 10.1016/j.devcel.2008.05.002

Li, J., Fan, S. L., Song, M. Z., Pang, C. Y., Wei, H. L., Li, W., et al. (2013). Cloning and characterization of a FLO/LFY ortholog in *Gossypium hirsutum* L. *Plant Cell Rep.* 32, 1675–1686. doi: 10.1007/s00299-013-1479-1

Li, J., Yuan, D., Wang, P., Wang, Q., Sun, M., Liu, Z., et al. (2021a). Cotton pan-genome retrieves the lost sequences and genes during domestication and selection. *Genome Biol.* 22:119. doi: 10.1186/s13059-021-02351-w

Li, L., Zhang, C., Huang, J., Liu, Q., Wei, H., Wang, H., et al. (2021b). Genomic analyses reveal the genetic basis of early maturity and identification of loci and candidate genes in upland cotton (*Gossypium hirsutum* L.). *Plant Biotechnol. J.* 19, 109–123. doi: 10.1111/pbi.13446

Liu, C., Xi, W., Shen, L., Tan, C., and Yu, H. (2009). Regulation of floral patterning by flowering time genes. *Dev. Cell* 16, 711–722. doi: 10.1016/j.devcel.2009.03.011

Liu, Z., Wu, X., Cheng, M., Xie, Z., Xiong, C., Zhang, S., et al. (2020). Identification and functional characterization of SOC1-like genes in *Pyrus bretschneideri*. *Genomics* 112, 1622–1632. doi: 10.1016/j.ygeno.2019.09.011

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

Ma, Z., He, S., Wang, X., Sun, J., Zhang, Y., Zhang, G., et al. (2018). Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fiber quality and yield. *Nat. Genet.* 50, 803–813. doi: 10.1038/s41588-018-0119-7

Ma, Z., Zhang, Y., Wu, L., Zhang, G., Sun, Z., Li, Z., et al. (2021). High-quality genome assembly and resequencing of modern cotton cultivars provide resources for crop improvement. *Nat. Genet.* 53, 1385–1391. doi: 10.1038/s41588-021-00910-2

Mouhu, K., Kurokura, T., Koskela, E. A., Albert, V. A., Elomaa, P., and Hytonen, 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

Nardeli, S. M., Artico, S., Aoyagi, G. M., De Moura, S. M., Da Franca Silva, T., Grossi-De-Sa, M. F., et al. (2018). Genome-wide analysis of the MADS-box gene family in polyploid cotton (*Gossypium hirsutum*) and in its diploid parental species (*Gossypium arboreum* and *Gossypium raimondii*). *Plant Physiol. Biochem.* 127, 169–184. doi: 10.1016/j.plaphy.2018.03.019

Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., et al. (2014). Structural basis for the oligomerization of the MADS domain transcription factor *SEPALLATA3* in *Arabidopsis*. *Plant Cell* 26, 3603–3615. doi: 10.1105/tpc.114.127910

Ren, Z., Yu, D., Yang, Z., Li, C., Qanmber, G., Li, Y., et al. (2017). Genome-wide identification of the MIKC-Type MADS-Box gene family in *Gossypium hirsutum* L. unravels their roles in flowering. *Front. Plant Sci.* 8:384. doi: 10.3389/fpls.2017.00384

Richter, R., Bastakis, E., and Schwechheimer, C. (2013). Cross-repressive interactions between SOC1 and the GATAs GNC and GNL/CGA1 in the control of greening, cold tolerance, and flowering time in *Arabidopsis*. *Plant Physiol.* 162, 1992–2004. doi: 10.1104/pp.113.219238

Rumpler, F., Theissen, G., and Melzer, R. (2018). A conserved leucine zipper-like motif accounts for strong tetramerization capabilities of *SEPALLATA-like* MADS-domain transcription factors. *J. Exp. Bot.* 69, 1943–1954. doi: 10.1093/jxb/ery063

Ruokolainen, S., Ng, Y. P., Albert, V. A., Elomaa, P., and Teeri, T. H. (2011). Over-expression of the Gerbera hybrida At-SOC1-like1 gene Gh-SOC1 leads to floral organ identity deterioration. *Ann. Bot.* 107, 1491–1499. doi: 10.1093/aob/mcr112

Ryu, C. H., Lee, S., Cho, L. H., Kim, S. L., Lee, Y. S., Choi, S. C., et al. (2009). *OsMADS50* and *OsMADS56* function antagonistically in regulating long day (LD)-dependent flowering in rice. *Plant Cell Environ.* 32, 1412–1427. doi: 10.1111/j.1365-3040.2009.02008.x

Sheldon, C. C., Conn, A. B., Dennis, E. S., and Peacock, W. J. (2002). Different regulatory regions are required for the vernalization-induced repression of *FLOWERING LOCUS C* and for the epigenetic maintenance of repression. *Plant Cell* 14, 2527–2537. doi: 10.1105/tpc.004564

Song, G. Q., Walworth, A., Zhao, D., Hildebrandt, B., and Leasia, M. (2013). Constitutive expression of the K-domain of a Vaccinium corymbosum SOC1-like (VcSOC1-K) MADS-box gene is sufficient to promote flowering in tobacco. *Plant Cell Rep.* 32, 1819–1826. doi: 10.1007/s00299-013-1495-1

Song, Q., Zhang, T., Stelly, D. M., and Chen, Z. J. (2017). Epigenomic and functional analyses reveal roles of epialleles in the loss of photoperiod sensitivity during domestication of allotetraploid cottons. *Genome Biol.* 18:99. doi: 10.1186/s13059-017-1229-8

Tadege, M., Sheldon, C. C., Helliwell, C. A., Upadhyaya, N. M., Dennis, E. S., and Peacock, W. J. (2003). Reciprocal control of flowering time by OsSOC1 in transgenic Arabidopsis and by FLC in transgenic rice. *Plant Biotechnol. J.* 1, 361–369. doi: 10.1046/j.1467-7652.2003.00034.x

Theissen, G., Melzer, R., and Rumpler, F. (2016). MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* 143, 3259–3271. doi: 10.1242/dev.134080

Turck, F., Fornara, F., and Coupland, G. (2008). Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annu. Rev. Plant Biol.* 59, 573–594. doi: 10.1146/annurev.arplant.59.032607.092755

Vekemans, D., Proost, S., Vanneste, K., Coenen, H., Viaene, T., Ruelens, P., et al. (2012). Gamma paleohexaploidy in the stem lineage of core eudicots: significance for MADS-box gene and species diversification. *Mol. Biol. Evol.* 29, 3793–3806. doi: 10.1093/molbev/mss183

Voogd, C., Wang, T., and Varkonyi-Gasic, E. (2015). Functional and expression analyses of kiwifruit SOC1-like genes suggest that they may not have a role in the transition to flowering but may affect the duration of dormancy. *J. Exp. Bot.* 66, 4699–4710. doi: 10.1093/jxb/erv234

Wang, S., Beruto, M., Xue, J., Zhu, F., Liu, C., Yan, Y., et al. (2015). Molecular cloning and potential function prediction of homologous SOC1 genes in tree peony. *Plant Cell Rep.* 34, 1459–1471. doi: 10.1007/s00299-015-1800-2

Wang, Z., Shen, Y., Yang, X., Pan, Q., Ma, G., Bao, M., et al. (2019). Overexpression of particular MADS-box transcription factors in heat-stressed plants induces chloroplast biogenesis in petals. *Plant Cell Environ.* 42, 1545–1560. doi: 10.1111/pce.13472

Wei, J., Liu, D., Liu, G., Tang, J., and Chen, Y. (2016). Molecular cloning, characterization, and expression of misoc1: a homolog of the flowering gene SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 from mango (*Mangifera indica* L). *Front. Plant Sci.* 7:1758. doi: 10.3389/fpls.2016.01758

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

Zhang, L., Wu, S., Chang, X., Wang, X., Zhao, Y., Xia, Y., et al. (2020). The ancient wave of polyploidization events in flowering plants and their facilitated adaptation to environmental stress. *Plant Cell Environ.* 43, 2847–2856. doi: 10.1111/pce.13898

Zhang, X., Wei, J., Fan, S., Song, M., Pang, C., Wei, H., et al. (2016). Functional characterization of GhSOC1 and GhMADS42 homologs from upland cotton (*Gossypium hirsutum* L.). *Plant Sci.* 242, 178–186. doi: 10.1016/j.plantsci.2015.05.001

Zhong, X., Dai, X., Xv, J., Wu, H., Liu, B., and Li, H. (2012). Cloning and expression analysis of GmGAL1, SOC1 homolog gene in soybean. *Mol. Biol. Rep.* 39, 6967–6974. doi: 10.1007/s11033-012-1524-0

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: 26 April 2022
doi: 10.3389/fpls.2022.861193

Characterization of *WRKY* Gene Family in Whole-Genome and Exploration of Flowering Improvement Genes in *Chrysanthemum lavandulifolium*

Muhammad Ayoub Khan†, Kang Dongru†, Wu Yifei, Wang Ying, Ai Penghui and Wang Zicheng*

State Key Laboratory of Crop Stress Adaptation and Improvement, Plant Germplasm Resources and Genetic Laboratory, Kaifeng Key Laboratory of Chrysanthemum Biology, School of Life Sciences, Henan University, Kaifeng, China

OPEN ACCESS

Edited by:
Yang Zhu,
Zhejiang University, China

Reviewed by:
Yifeng Xu,
Nanjing Agricultural University, China
Yong Zhuang,
Jiangsu Academy of Agricultural Sciences (JAAS), China
Yichao Li,
St. Jude Children's Research Hospital, United States
Zhong Qi Fan,
Fujian Agriculture and Forestry University, China

***Correspondence:**
Wang Zicheng
ZCWang@henu.edu.cn

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

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

Received: *24 January 2022*
Accepted: *02 March 2022*
Published: *26 April 2022*

Citation:
Ayoub Khan M, Dongru K, Yifei W, Ying W, Penghui A and Zicheng W (2022) Characterization of WRKY Gene Family in Whole-Genome and Exploration of Flowering Improvement Genes in Chrysanthemum lavandulifolium. Front. Plant Sci. 13:861193. doi: 10.3389/fpls.2022.861193

Chrysanthemum is a well-known ornamental plant with numerous uses. *WRKY* is a large family of transcription factors known for a variety of functions ranging from stress resistance to plant growth and development. Due to the limited research on the *WRKY* family in chrysanthemums, we examined them for the first time in *Chrysanthemum lavandulifolium*. A total of 138 *ClWRKY* genes were identified, which were classified into three groups. Group III in *C. lavandulifolium* contains 53 members, which is larger than group III of *Arabidopsis*. The number of introns varied from one to nine in the *ClWRKY* gene family. The "WRKYGQK" motif is conserved in 118 members, while other members showed slight variations. AuR and GRE responsive *cis*-acting elements were located in the promoter region of *WRKY* members, which are important for plant development and flowering induction. In addition, the W box was present in most genes; the recognition site for the *WRKY* gene may play a role in autoregulation and cross-regulation. The expression of the most variable 19 genes in terms of different parameters was observed at different stages. Among them, 10 genes were selected due to the presence of CpG islands, while nine genes were selected based on their close association with important *Arabidopsis* genes related to floral traits. *ClWRKY36* and *ClWRKY45* exhibit differential expression at flowering stages in the capitulum, while methylation is detected in three genes, including *ClWRKY31, ClWRKY100,* and *ClWRKY129.* Our results provide a basis for further exploration of *WRKY* members to find their functions in plant growth and development, especially in flowering traits.

Keywords: *WRKY* transcription factors, ornamental, motifs, *cis*-acting elements, CpG islands, methylation, flowering traits

INTRODUCTION

Chrysanthemum lavandulifolium is one of the original species of *Chrysanthemum* × *morifolium* and uses as a model plant. It is a major ornamental plant, originated in China, an important cut flower with diversified petal color and shape having a variety of flower structures, which makes it a highly valuable ornamental plant in the floral industry worldwide (Nguyen and Lim, 2019). Chrysanthemum is among the four most important ornamental plants in the world used for

beautification and aesthetic value (Shinoyama et al., 2012). The flower of chrysanthemum is also known for its health benefits, and it contains important antioxidants and phenolic compounds (Lin and Harnly, 2010; Yue et al., 2018).

Due to a simple diploid ($2n = 2x = 18$) genetic background, *C. lavandulifolium* dominates over *C. morifolium* (Huang et al., 2012; Dong-ru et al., 2019). The wild ancestors having desired qualities against biotic and abiotic cues can be used for the genetic improvement of cultivated varieties. Identification and incorporation of such genes into commercial varieties reduces pesticide use, improves plant health, and induces photoperiod-controlling flowering mechanisms (Wu et al., 2010; Fu et al., 2013; Yang et al., 2017).

Genetic architectural study is one of the major research directions in any plant. By narrowing the path further, transcription activation and silencing is the major field of concern. Many transcription factors (TFs) have been explored to switch on and off against certain environmental conditions. *WRKY* TF is one of the major families first identified in plants in 1994 by the name of DNA-binding protein in sweet potato, which is labeled as *SWEET POTATO FACTOR1 (SPF1)* (Ishiguro and Nakamura, 1994). *WRKY* TFs play important roles in resistance against biotic (fungal or bacterial pathogen) and abiotic stresses, growth, and development. Several *WRKY* TFs have been proven to provide resistance against biotic stresses exerted by fungal or bacterial pathogens by influencing other associated genes (Turck et al., 2004; Duan et al., 2007; Zhou et al., 2008, 2011; Li et al., 2009, 2011; Chen et al., 2010; Scarpeci et al., 2013; Phukan et al., 2016; Yokotani et al., 2018; Warmerdam et al., 2020). The name *WRKY* has been given to this family for their distinguishing characteristic, the *WRKY* domain, which is composed of highly conserved 60 amino acids and a zinc finger motif. *WRKY* TFs are classified into three groups based on the number of (WRKYGQK) domain, and zinc-finger motif. Group I has two domains, and groups II and III have one domain. Groups I and II have C2H2 type zinc finger motifs while group III has a C_2HC zinc finger motif. Members of group II were further classified into subgroups (Rushton et al., 2010).

WRKY TFs are extensively studied in *Arabidopsis* and most other plants. *WRKY7* and *OsWRKY11* expression resulted in blooming delay in *Arabidopsis* and rice (Cai et al., 2014; Chen et al., 2019), *WRKY13* has an adverse effect on flowering time, and *WRKY*12, *GsWRKY20* and *CsWRKY50* promote flowering, possibly through the influence of gibberellic acid under short-day conditions in *Arabidopsis* (Luo et al., 2013; Kumar et al., 2016; Li et al., 2016), *GhWRKY22* regulating pollen development (Wang et al., 2019; Ramos et al., 2021). While different *WRKY* family TFs play an important role in chrysanthemum, they showed resistance to salt stress (Liu et al., 2013, 2014; Li et al., 2015b; Jaffar et al., 2016; Liang et al., 2017; Wang K. et al., 2017; He et al., 2018; Zhang et al., 2020), and *CmWRKY1* and *CmWRKY15* transgenic lines in *C. morifolium* interact with *SA* and *ABA* and altered plant development (Fan et al., 2015, 2016; Bi et al., 2021). *CmWRKY48* showed resistance against aphid infestation (Li et al., 2015a). They also play a regulatory role in the flowering stages. The complex network of *VQ20*, *WRKY2*, and *WRKY34* interaction influences the plant gametogenesis and plays a crucial part in the development of pollen and pollen tube growth, modulating flowering (Lei et al., 2017; Ma et al., 2020).

Flowering time is a complex character regulated by a network of different external and internal signals. Different pathways have been identified, including photoperiod pathway, vernalization pathway, gibberellin-regulated pathways, and autonomous floral initiation pathway (Soares et al., 2020). Therefore, flowering induction studies are more crucial in ornamental plants, especially in chrysanthemum, where all the ornamental characteristics are associated with flowering, i.e., flower color, shape, and size. Chrysanthemum is a short-day plant and needs a specific photoperiod to induce flowering, which makes it relatively expensive as well as limits the annual production of chrysanthemum flowers. We explored the *WRKY* gene family in *C. lavandulifolium* for the first time, prior, the *WRKY* gene family has only been studied in the *C. morifolium* (Song A. et al., 2014). The size of the *WRKY* family in *C. morifolium* comprised 15 members, which were named from *CmWRKY1* to *CmWRKY15* (GenBank: KC615355–KC615369). As *WRKY* is a large family, influencing floral characteristics, we need to explore them deeply. Our present study was a part of such investigation in which we tried to explore the *WRKY* gene family in *C. lavandulifolium* comprised structural gene analysis, protein motifs, phylogenetic relatedness of *C. lavandulifolium* and *Arabidopsis thaliana*, *cis*-acting elements, recognition of CpG rich regions, the methylation status of the genes, and expression of 19 important genes in different tissues at different stages to check their role in flowering-related traits. This investigation will lead us to identify variation in transcription regions in the *WRKY* family for future chrysanthemum research.

MATERIALS AND METHODS

Plant Material and Treatment

The plant materials of *C. lavandulifolium* were obtained from Beijing Forestry University. The plants' G1 lines were grown under controlled environmental conditions in the growth chamber. All the plants used in this research came from the G1 line. The plants were propagated by cuttings and grown in pots. The plants were kept under long-day conditions (14 h L/10 h D) light/dark at a 25°C temperature. The photoperiod conditions were changed to short-day conditions (12 h L/12 h D) when the plants attained maturity (more than 14 leaves). The temperature was kept constant. The plants were watered every 7 days. Stem, root, leaf, buds, and capitulum samples for quantitative real-time polymerase chain reaction (qRT-PCR) were taken at different growth stages, including young, mature, and flowering. The plants were considered at a young stage when they attained eight leaves. The plants were called into the adult stage when the number of leaves was 14, and they were transferred to short-day conditions. The flower develops in short conditions, and samples were taken at different stages of flower development (**Supplementary Figure 1**). The little buds (1–3 mm in diameter) appeared after 7 days in short-day conditions and were taken for RNA extraction. The median buds (5–8 mm in diameter) were

taken after a few days before flower opening. The first flower samples were taken when the flowers had just opened.

Characterization of *WRKY* Gene Family in *Chrysanthemum lavandulifolium*

The *C. lavandulifolium* genome has been sequenced (Wen et al., 2022). All the protein and genomic sequences of *ClWRKY* members were obtained from PRJNA681093, National Center for Biotechnology Information (NCBI). The genomic information of *C. lavandulifolium* was completed by Henan University and Beijing Forestry University, and related papers have been published (Wen et al., 2022). The redundant sequences, which lack the *WRKY* domain, were removed after the sequence analysis using manual inspection in the Mega X software. A total of 138 definite *ClWRKY* sequences were taken for further study. The *ClWRKY* members were named from *ClWRKY1* to *ClWRKY138*, according to the source provider (**Supplementary Table 3**).

Classification of *ClWRKY* Proteins

Multiple protein sequence alignments of *C. lavandulifolium* and *A. thaliana WRKY* were made using ClustalX and MEGA-X version 10.2.6 (Molecular Evolutionary Genetics Analysis) (Kumar et al., 2018). The phylogenetic tree based on amino acid (aa) sequences of *ClWRKY* and *AtWRKY* conserved *WRKY* domains was constructed following the neighbor-joining method with 1,000 bootstraps to determine the close association of both gene families.

Gene Structure and Protein Motif Analysis

The gene structure (intron-exon) of *ClWRKY* members, full-length gene, and coding (CDS) sequences were analyzed in the online database, Gene Structure Display Server (GSDS)[1] as described (Hu et al., 2015). Analysis of the conserved protein motifs was performed using the NCBI CDD batch[2] and the TBtools software (Chen et al., 2020).

Cis-Acting Elements in the Promoter Region of *ClWRKY* Members

The promoter sequences of all 138 *ClWRKY* members were submitted to the PlantCARE software to analyze their promoter region for *cis*-acting elements (Lescot et al., 2002). For this purpose, a 2,000 bp sequence upstream of the gene was taken to identify *cis*-acting elements. The MEME suite software was used to identify *de novo* motifs and known enriched motifs (Bailey et al., 2015). The *de novo* motif analysis was performed to identify significant patterns (Bailey and Elkan, 1994).[3] The known enriched motifs were identified *via* simple enrichment analysis using the MEME software (Bailey and Grant, 2021).[4] Tomtom motif comparison tool was used to quantify the similarity between motifs (Gupta et al., 2007).

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

[2]https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi

[3]https://meme-suite.org/meme/tools/meme

[4]https://meme-suite.org/meme/tools/sea

CpG Island Analysis of the Promoter Region of *ClWRKY* Gene Family

For CpG analysis, promoter regions from the *WRKY* gene family were sequenced. For the exploration of CpG islands, the "Methyl Primer Express version 1.0" software was used (Methyl Primer Express Software v1.0 Quick Reference Card).[5] All the sequences were submitted to the software one by one to identify CpG islands or rich regions. The minimum length of the island was kept at 200 bp, while the maximum length was 2,000 bp. The criterion of "C + G/Total bases" was 50%. The CpG observed/CpG expected ratio was 0.6.

Genomic DNA Extraction and Bisulfite Treatment

The total DNA was extracted from different tissues of *C. lavandulifolium*. The tissues included root, stem, leaf, and capitulum at three different growth stages, i.e., young stage, adult stage, and flowering stage. Total genomic DNA was extracted using the "Plant Genomic DNA Kit" (TOLOBIO), according to the manufacturer's instructions. DNA was quantified using Thermo Scientific NanoDrop^{TM} 2000/2000c Spectrophotometers, and the quality of DNA was checked by 1.5% agarose gel electrophoresis.

Total genomic DNA from each sample was treated with sodium bisulfites using the "DNA Bisulfite Conversion Kit" (TIANGEN) according to the manufacturer's instructions. The conversion efficiency and quality were measured using Thermo Scientific NanoDrop^{TM} 2000/2000c Spectrophotometers.

Methylation-Specific Polymerase Chain Reaction

Ten genes were selected on the basis of CpG analysis. Primers for all ten genes were designed using the "Methyl primer express version 1.0" software. For this purpose, two types of primers (methylated and unmethylated) were designed to check the possible methylated sites. The primer length ranged from 18 to 22 bp. The PCR amplicon length was between 100 and 175 bp. Methylation-Specific Polymerase Chain Reaction (MSP) was conducted to analyze the CpG island among the ten selected genes after treatment with bisulfite for methylation status. The MSP conditions were as follows: pre-denaturation for 3 min at 95°C, followed by 34 cycles of the 30s at 95°C and 30 s at 55°C, 1 min at 72°C, and a final extension for 5 min at 72°C. PCR products were resolved onto 1.5% agarose gel and observed under UV transilluminator.

Quantitative Real-Time Polymerase Chain Reaction

Nineteen (Zhou et al., 2011) variable genes on the basis of CpG island and phylogenetic analysis were selected for gene expression patterns. Primers for all the 19 *ClWRKY* genes were designed using the "Primer Premier 5" software. RNA was extracted from different plant tissues, i.e., root, stem, and leaves at young, adult, and flowering stages, while the capitulum samples, which

[5]thermofisher.com

comprise little buds, median buds, and first flower (right at the time of flower opening), were taken at flowering stages. The RNA was extracted using the RNAprep Pure Plant Plus Kit (spin column) according to the manufacturer's instructions. The cDNA was synthesized using the "StarScript II First-strand cDNA Synthesis Mix with gDNA Remover." The expression of all the selected genes was observed in qRT-PCR. The qRT-PCR was carried out on light cycle "Quantageneq225 Fluorescent Quantitative PCR," using "2 × RealStar Green Fast Mixture." The conditions of PCR were as follows: denaturation at 95°C for 2 min, followed by 40 cycles of denaturation at 95°C for 15 s, annealing at 60°C for 20 s, and extension at 65°C for 5 s. The actin gene of the chrysanthemum was used as an internal control. The data from qRT-PCR for relative expression were analyzed using the $2^{-\Delta\Delta Ct}$ method proposed by Livak and Schmittgen (2001).

RESULTS

Identification and Characterization of *WRKY* Gene Family in *Chrysanthemum lavandulifolium*

WRKY genes possess a conserved "WRKYGQK" domain. The "WRKYGQK" heptapeptide is the most defining characteristic of *WRKY* TFs. The *C. lavandulifolium WRKY* sequences were obtained from PRJNA681093, NCBI. A total of 153 *WRKY* gene members have been selected; among them, 138 were exhibited for sequence analysis, and 15 were considered defective. The *WRKY* members are named from *ClWRKY*1 to *ClWRKY*138 according to their location on the chromosome. Protein motif analysis showed homology in most of the conserved regions. Among 138 *ClWRKY* members, 85.51% of members have homology in the "WRKYGQK" core domain, and 19 members have two conserved *WRKY* domains. However, 14.5% of *WRKY* members showed variation in the conserved heptapeptide domain and had substitutions in the core domain. Out of 14.5% variants, 13 members (*ClWRKY*37, 38, 39, 49, 50, 51, 52, 53, 54, 74, 76, 77, and 98) have "WRKYGKK," substituting glutamine as a lysine aa. *ClWRKY*12 replaced histidine instead of lysine, two genes, *ClWRKY*20 and *ClWRKY21*, shared sixth and seventh as glutamine and asparagine substitutions. High diversity has been found in three genes, i.e., (*ClWRKY*27) "WKKYGEQK," (*ClWRKY*97) "WKKYGEKK," and (*ClWRKY*131) "WRKNGQN" *WRKY* domains, respectively (**Figure 1**).

For the determination of important properties such as aa length, molecular weight, and isoelectric point, we thoroughly analyzed the *WRKY* gene. The average length of *WRKY* proteins was 313 aa, while 74 aa was the minimum and 697 aa was the maximum length. The *WRKY* members with the least and highest lengths were *ClWRKY58* and *ClWRKY36*, respectively. The molecular weight ranged from 0.02 kDa (*ClWRKY55*) to 76.26 kDa (*ClWRKY36*). The pIs (isoelectric point) ranged from 4.73 (*ClWRKY75*) to 10.63 (*ClWRKY97*). Among 138 *ClWRKY* members, 82 were acidic (having less than 7 pI), while 56 were basic (having more than 7 pI). More details of *ClWRKY* features are available in **Supplementary Table 1**.

Phylogenetic Analysis and Classification of *ClWRKY* Proteins

A phylogenetic tree was constructed between *C. lavandulifolium* and *A. thaliana* to investigate the relatedness of different *WRKY* genes and to understand the classification of different groups among *WRKY* members. *Arabidopsis* is a model plant for genetic study and has been explored for the function of *WRKY* genes. The MEGA-X software was used for agglomerative construction following the neighbor-joining method to find the correlation between *WRKY* members of *C. lavandulifolium* and *A. thaliana*. The *WRKY* members are divided into three groups or seven subgroups. Group I has 27 members, and group II has 58 members with 9 members in IIa, 10 members in IIb, 10 members in IIc, 10 members in IId, and 19 members in IIe. Group III has 53 members and was found to be the largest group in this classification (**Figure 2**). The majority of the *ClWRKY* members clustered in the corresponding groups. Some of the *WRKY* members (*ClWRKY20*, *21, 55, 56,* and *58*) had one *WRKY* domain, but still, clustered in group I; by the rules, group I members should have two domains. These *WRKY* members may evolve from two *WRKY* domain genes and lose one domain in the process of evolution. Previously, such exceptional cases of *WRKY* members with one *WRKY* domain clustering in group I have been reported (Bi et al., 2016).

Some of the *ClWRKY* members clustered with important *Arabidopsis WRKY* genes that helped us identify the genes that are associated with flowering. *ClWRKY36* is clustered with *Arabidopsis AtWRKY2* and *AtWRKY34* (**Figure 2**), both of which play an important role in pollen development through interaction with the *VQ20* protein (Lei et al., 2017). Similarly, *ClWRKY45* clustered with *AtWRKY71* (**Figure 2**), which play a key role in the regulation of flowering, directly and indirectly, by influencing other flowering-associated genes (Yu et al., 2016, 2018).

Gene Structure of *WRKY* Members and Protein Motif Analysis

For the determination of intron-exon structure similarity, Gene Structure Display Server (GSDS)[6] database has been used. Variation among *ClWRKY* members was observed in terms of a number of introns. The minimal number of introns in a gene was one, which occurred in seven members including *ClWRKY27, ClWRKY58, ClWRKY59, ClWRKY76, ClWRKY86, ClWRKY92*, and *ClWRKY125* while *ClWRKY90* had 9 introns, which is the maximal in the whole family. The overall average number of introns in the *ClWRKY* family was three (**Figure 3**). The number of introns in group I ranged from 1 to 5, group IIa varied from 3 to 9, group IIb from 2 to 6, group IIc from 1 to 2, group IId from 1 to 4, and group IIe from 1 to 4. Group III varied from 1 to 6 (**Supplementary Table 2**). The minimal number of exons was two, which appeared in *ClWRKY27, ClWRKY58, ClWRKY59, ClWRKY76, ClWRKY86, ClWRKY92,*

[6]http://gsds.cbi.pku.edu.cn/

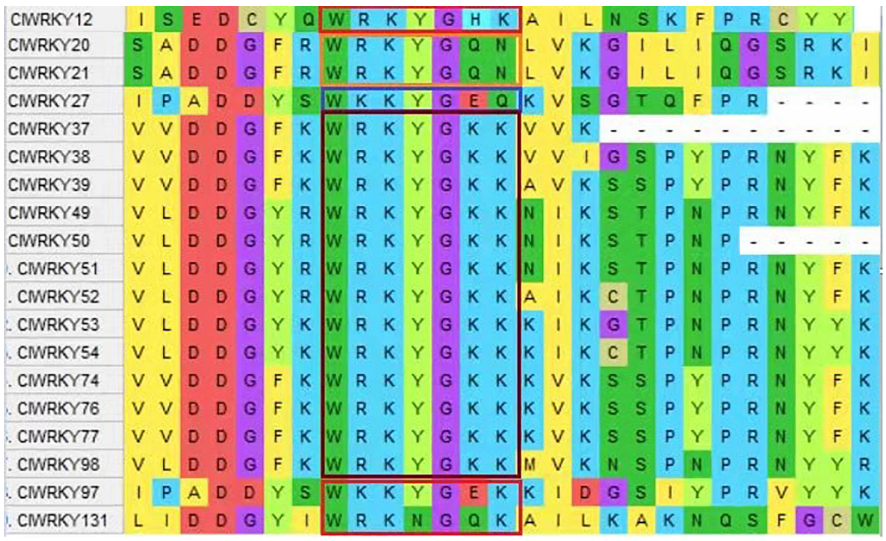


FIGURE 1 | A screenshot of the MEGA X version 10.2.6, showing variation in the conserved domain (WRKYGQK) of *ClWRKY* proteins, and the variation of the conserved domain is encircled in different colors.

and *ClWRKY125,* while the maximal number of exons was 10, which occurred in *ClWRKY87* and *ClWRKY90* (**Supplementary Table 2**). The overall average number of exons in the *ClWRKY* family was four (**Figure 3**). The exon numbers varied within the groups; group I ranged from 2 to 6, group IIa from 4 to 10, group IIb from 3 to 7, group IIc from 2 to 3, group IId from 2 to 5, group IIe from 2 to 4, and group III from 2 to 7, respectively. The average number of exons within the groups is as follows: group I has 4, group IIa has 6, group IIb has 5, group IIc has 3, group IId has 3, group IIe has 3, and group III has 3, respectively (**Supplementary Table 2**). Protein motif analyses were performed on the *ClWRKY* gene family. The first two motifs represented by green and yellow colors are *WRKY* protein motifs (**Figure 4**). At least one or two *WRKY* protein motifs have been identified among all the gene sequences. This suggests that *WRKY* is a conserved domain.

CpG Island Analysis of the Promoter Region of *ClWRKY* Gene Family

CpG island is the region of DNA that is rich in CpG sites and is often located near the promoters of genes. Ten members were found to have CpG islands; among them, seven members contained CpG in their promoter region, two in the cross-promoter and gene region, and one member was found to have it in the gene body. All the ten members had one CpG island. *ClWRKY115* had the shortest island with 438 bp, while *ClWRKY83* had the longest island with a length of 1,065 bp (**Table 1**). The genes with the CG-rich content are the obvious means of DNA methylation and a prominent source of epigenetic improvement.

Identification of DNA Methylation

Total genomic DNA of *C. lavandulifolium* was extracted from the root, stem, leaf, little buds, median buds, and first flowers (flowers at just opening time). These tissues were taken at different growth stages of the plant, which include the young stage, adult stage, and flowering stage. The methylation and unmethylation primers were designed for the ten different genes using the "Methyl Primer Express version 1.0" software. These genes were selected due to the presence of CpG islands. The MSP was carried out for the following ten genes: *ClWRKY9, ClWRKY31, ClWRKY33, ClWRKY41, ClWRKY83, ClWRKY97, ClWRKY100, ClWRKY115, ClWRKY129*, and *ClWRKY130*. The methylation status detected by MSP was analyzed through gel electrophoresis. The MSP results showed the presence of methylation in *ClWRKY31, ClWRKY100*, and *ClWRKY129* at different stages in different tissues. These genes showed significant methylation status in all tissue samples at all growth stages of the plant (**Figures 5A–C**). However, the methylation status detection through bisulfite sequencing can give us more clear results compared with the MSP method.

Cis-Acting Elements in the Promoter Region of *ClWRKY*

The *cis*-acting elements in the promoter region are really important for gene expression (Rombauts et al., 1999). The 2,000 bp sequence upstream of the genes was submitted to the PlantCARE software and searched for *cis*-acting elements. A range of *cis*-acting elements was identified in the promoter region of the Cl*WRKY* family, which were highly diverse.

A total of 110 *cis*-acting elements were found in the promoter region of *ClWRKY* members. Most of these elements have a diverse role in the plant life cycle. According to our objectives, we screened out the elements that can play a role in plant growth and development and particularly in flowering induction. Therefore, we selected 15 *cis*-acting elements on the basis of their possible role in flowering-related traits and the overall development of

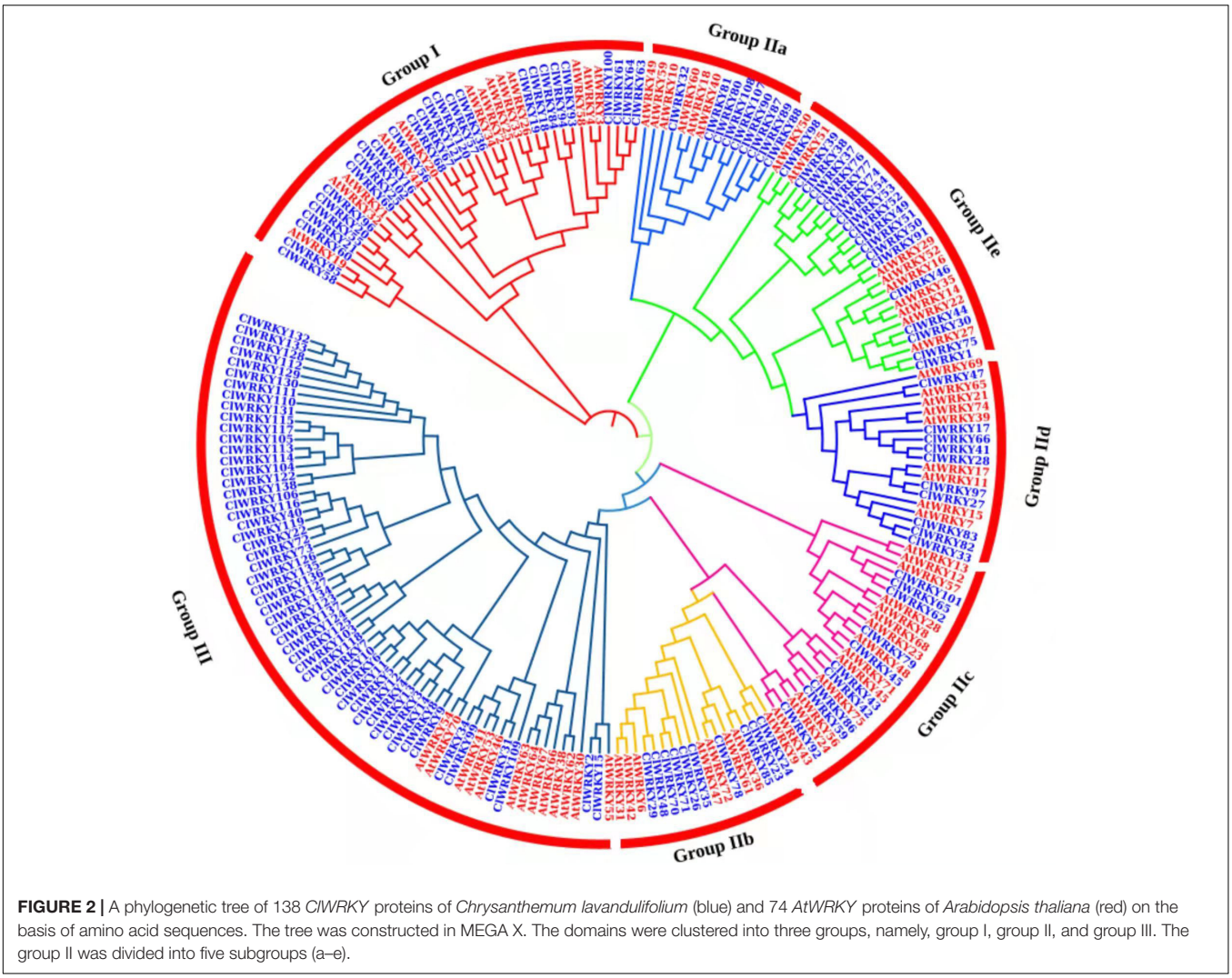


FIGURE 2 | A phylogenetic tree of 138 *ClWRKY* proteins of *Chrysanthemum lavandulifolium* (blue) and 74 *AtWRKY* proteins of *Arabidopsis thaliana* (red) on the basis of amino acid sequences. The tree was constructed in MEGA X. The domains were clustered into three groups, namely, group I, group II, and group III. The group II was divided into five subgroups (a–e).

the plant. The detailed features of the 15 *cis*-acting elements are provided in **Supplementary Table 4**.

Most of the genes have important *cis*-acting elements, such as auxin responsiveness (AuR) and gibberellin responsive elements (GRE), associated with flowering-related traits (**Figure 6**). W-box was most widely available among the *ClWRKY* family. *ClWRKY45* and *ClWRKY36*, the genes that resulted in high expression at the capitulum stage, had some important elements involved in endosperm expression, elements involved in auxin responsiveness, and ethylene-responsive elements (AuR, EE, and ERE), which are supposed to play a role in flower development. Methyl jasmonate (MeJAR), which plays important role in flowering induction, was present in most of the *ClWRKY* members, especially in group III members, which indicates that *WRKY* members may play a role in the overall development of a plant.

The *de novo* motif analysis in MEME identified 15 motifs (**Supplementary Figure 2**). Tomtom tool comparison showed homology with some important motifs, such as motifs similar to MEME12 and MEME14, which are involved in seed germination, pollen tube growth, and WRKY binding sites (**Supplementary Table 5**). The enriched motifs identified in MEME are mostly different from the motifs found in the PlantCare software. The simple enrichment analysis (SEA) in MEME identified 95 known enriched motifs. Among them, 49 motifs encoded ethylene, which is essential for plant growth and development. Other important motifs included are WRKY binding sites, TEOSINTE BRANCHED, cycloidea and PCF (TCP) binding sites, and so on (**Supplementary Table 6**).

qRT-PCR Analyses

Different *ClWRKY* genes having special characters were selected on the basis of previous analysis. Total RNA was extracted from different tissues (root, stem, leaves, and flowers) *of C. lavandulifolium* at different stages (young, adult, and flowering). The expression pattern of 19 *WRKY* genes was observed in all tissues at different stages of the plant's life span (**Figure 7**). Among them, 10 genes were selected due to the presence of CG-rich regions, and nine genes were selected as they showed close association with important *Arabidopsis WRKY* genes that are associated with floral traits. Most of the genes exhibited higher expression at the flowering stage in the capitulum

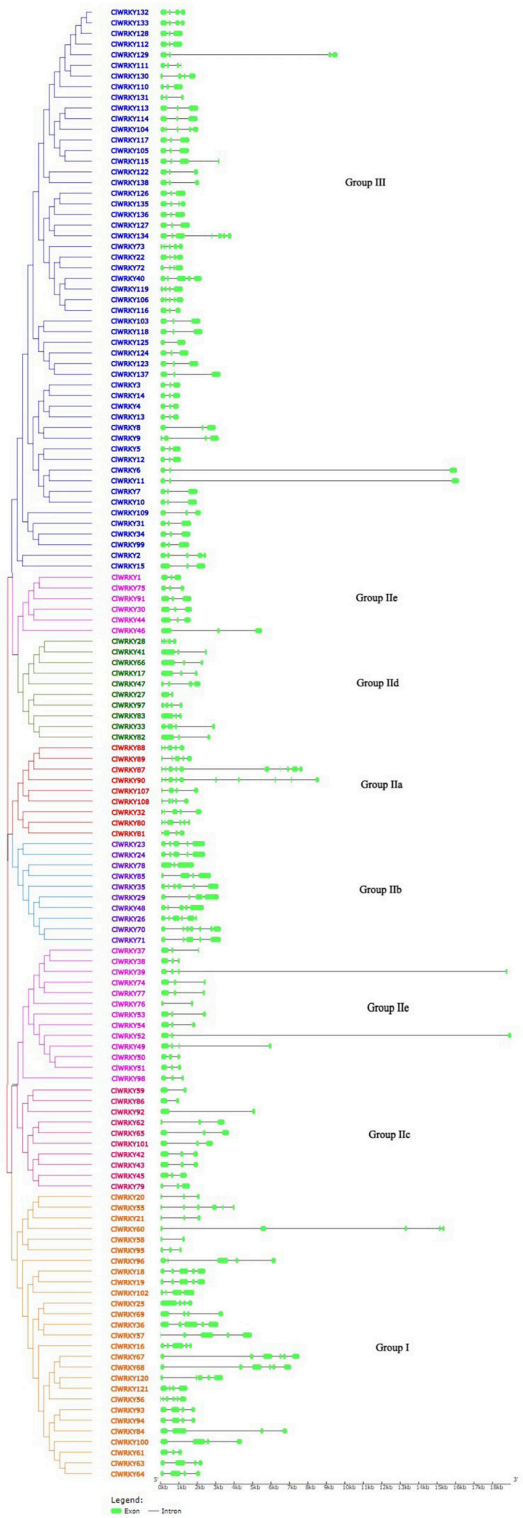


FIGURE 3 | The intron-exon structure of 138 *ClWRKY* members according to the phylogenetic relationship. A phylogenetic tree was constructed with the *ClWRKY* protein sequences using MEGA X. The Gene Structure Display Server (GSDS) software was used to determine the intron-exon structure. Green bars represent exons in the gene. Different colors of *WRKY* members represent different groups in the *ClWRKY* family.

tissues. The highest expression in the capitulum tissues at the flowering stage was exhibited by *ClWRKY36*, *ClWRKY45, ClWRKY57, ClWRKY59,* and *ClWRKY82* (**Figure 7**). The expression of these genes was higher in the flowering stage, especially in the first flowering (FF) stage and median bud stage, which suggests their role in flowering regulation. These genes were selected for qRT-PCR as they clustered with flowering-related At*WRKY* genes (**Figure 2**). On the basis of previous analysis, *ClWRKY36* clustered with *ClWRKY71*, so it might have a basic role in the control of flowering time and plant development, as *WRKY71* homolog was found to promote flowering in woodland strawberries (Lei et al., 2020). *ClWRKY31*, *ClWRKY33, ClWRKY41, ClWRKY97, ClWRKY100, ClWRKY129,* and *ClWRKY130,* which were selected on the basis of having CpG rich regions (**Supplementary Table 1**), showed higher expression in capitulum tissues (**Figure 7**). In addition, *ClWRKY31, ClWRKY100,* and *ClWRKY129* were detected with DNA methylation (**Figure 5**), which is linked with flowering regulation and is one of the basic sources of epigenetics. Interestingly, *ClWRKY83* remains stable across all the tissues and stages as it is constantly expressed in all samples of *C. lavandulifolium*. It might play a role in adaptability. Moreover, *ClWRKY83* showed close association with important *AtWRKY7* of *Arabidopsis*. The other *WRKY* member's expressions showed less or non-significant changes at different stages in all tissues.

We also investigated *ClLFY* (leafy), as it not only promotes flowering time but also influences other genes that regulate flowering-related traits (Yang et al., 2016). The expression of this gene was high in the median bud and adult leaf stages, but low in stem and root (**Figure 7**). Besides, *ClLFY* may promote flowering time in *C. lavandulifolium* by making association with *ClWRKY36 and ClWRKY45*.

DISCUSSION

WRKY TFs are one of the largest transcription families, playing an important role in plant development and resistance against different biotic and abiotic stresses. *WRKY* TFs are spread broadly across higher plants (Li et al., 2014). *WRKY* family investigation is limited to *C. morifolium* with 15 genes (Song A. et al., 2014). Being a model plant for chrysanthemums (Dongru et al., 2019), the characterization of the *WRKY* gene family in *C. lavandulifolium* is of great importance. For the first time, we identified 138 *WRKY* members in *C. lavandulifolium*. Out of 138 members, 118 were properly conserved in the "WRKYGQK" domain. Substitution of glutamine occurs instead of lysine in thirteen *WRKY* heptapeptide domains. This WRKYGKK sequence is reported to be a major variant in many studies (Song H. et al., 2014; Song et al., 2016a,b). *ClWRKY12* (WRKYGHK) "Q" replaced by "H" *ClWYKY20* and 21 (WRKYGQN) "K" replaced by "N," *ClWRKY27* (WKKYGEQK) "K" replaced by "Q," *ClWRKY*97 (WKKYGEK) "Q" replaced by "E," and *ClWRKY131* (WRKNGQK) "Y" replaced by "N" as their respective domains were also reported in carrot (Nan and Gao, 2019). The variation in the heptapeptide domain of *WRKY* genes

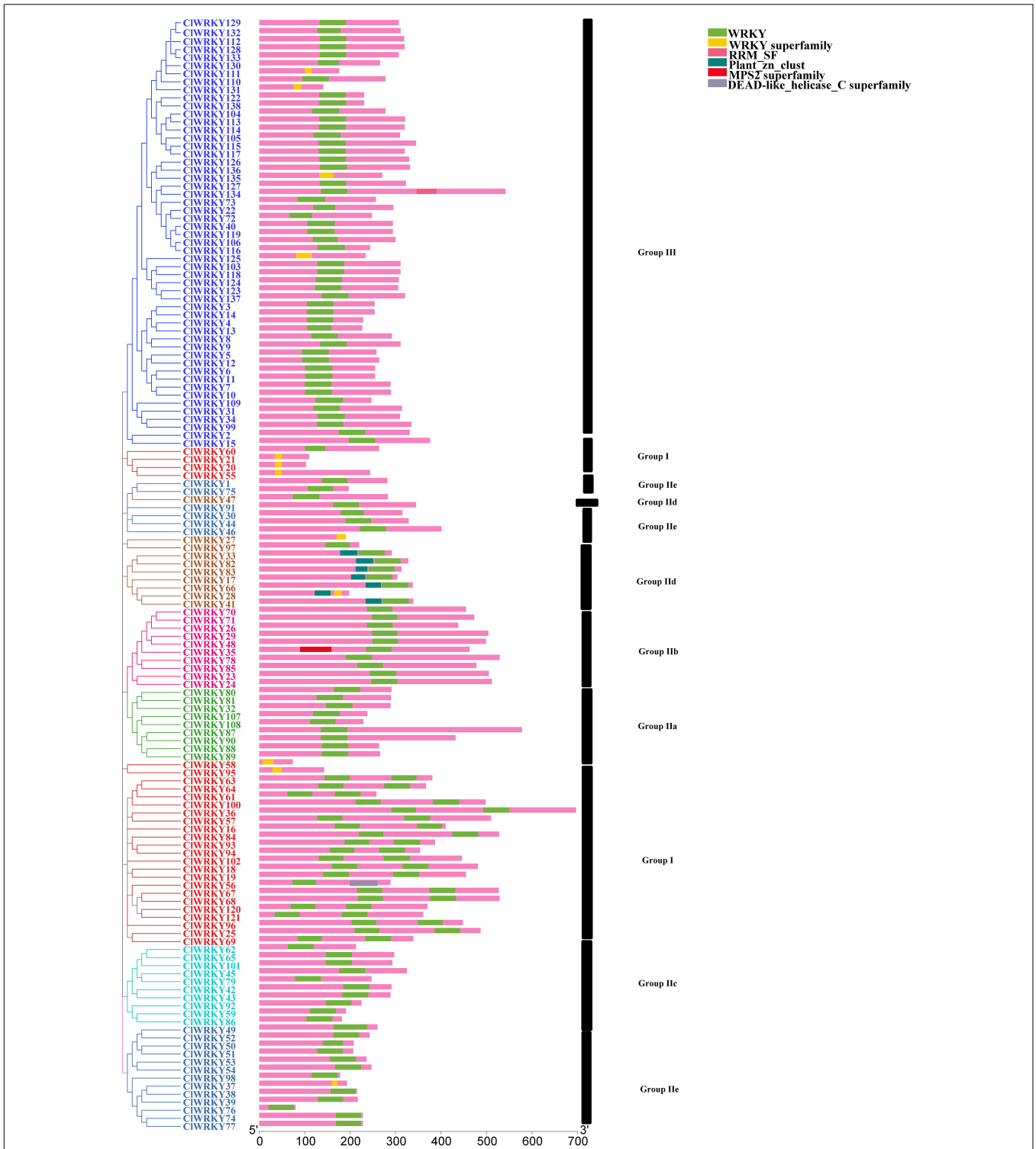


FIGURE 4 | Schematic diagram of conserved motifs in *ClWRKY* proteins according to the phylogenetic relationship. Different color *ClWRKY* members indicate different groups. The conserved motifs were predicted using the TBtools software.

has been largely demonstrated (Yamasaki et al., 2005; Brand et al., 2013; Rinerson et al., 2015; Yang et al., 2020); however, variations in aa sequence were different from our findings. The variation in the WRKYGQK conserved domain mainly occurs from Q to K aa. The same kind of substitution was observed in the *WRKY* gene family of *Camelina sativa* (Song et al., 2020).

TABLE 1 | Details of each of the ten genes with a CpG island.

NO	*WRKY* Gene	Island #	Start	End	Length	Region
1	*ClWRKY9*	1	–1239	–593	647	Promoter
2	*ClWRKY31*	1	–1049	–251	799	Promoter
3	*ClWRKY33*	1	–4	566	571	Promoter + Gene Body
4	*ClWRKY41*	1	–1296	–688	609	Promoter
5	*ClWRKY83*	1	–2000	–935	1,065	Promoter
6	*ClWRKY97*	1	–80	471	552	Promoter + Gene Body
7	*ClWRKY100*	1	–2000	–1429	571	Promoter
8	*ClWRKY115*	1	–2000	–1562	438	Promoter
9	*ClWRKY129*	1	6494	6950	457	Gene Body
10	*ClWRKY130*	1	–878	–251	628	Promoter

The 138 ClWRKY genes were analyzed for CG-rich regions using the "Methyl Primer Express version 1.0" software.

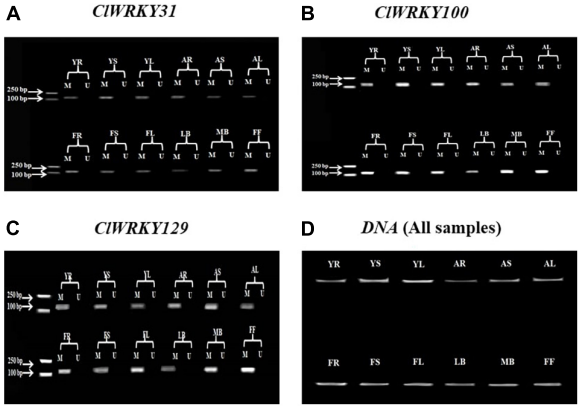


FIGURE 5 | *ClWRKY31, ClWRKY100*, and *ClWRKY129* Methylation-Specific Polymerase Chain Reaction (MSP) results **(A–C)**. An MSP was conducted to determine the methylation status of selected genes. The PCR product length of the genes was as follows: *ClWRKY31* = 156 bp, *ClWRKY100* = 135 bp, and *ClWRKY129* = 131 bp. "M" represents a methylated PCR product, and "U" represents an unmethylated PCR product. **(D)** The genomic DNA results for all samples. These results were verified through gel (1.5%) electrophoresis. The abbreviations of different growth stages and tissues are as follows: YR, young stage root; YS, young stage stem; YL, young stage leaf; AR, adult stage root; AS, adult stage stem; AL, adult stage leaf; FR, flowering stage root; FS, flowering stage stem; FL, flowering stage leaf; LB, little buds; MB, median buds; FF, first flower at just opening time.

The *ClWRKY* gene family is composed of 138 genes, which is bigger than the *AtWRKY* gene family (Rushton et al., 2010). The *ClWRKY* family was classified into three groups. The classification of *ClWRKY* members was according to the classification of the *WRKY* family in *Arabidopsis*. Most of the *ClWRKY* members clustered in the corresponding groups. However, some of the *ClWRKY* members with one *WRKY* domain clustered in group I, which included, *ClWRKY58, 95, 60, 56, 20, 21,* and *ClWRKY55*. Among them, *ClWRKY20*, *21,* and *55* possessed a unique pattern "WNSIVV" which was conserved in these three members. Clustering of one domain *WRKY* of factors in group I is consistent with the study of the *WRKY* TF family in *Linum usitatissimum* (Yuan et al., 2021). The same results are reported in *Cicer arietinum* (Kumar et al., 2016). Some *WRKY* members with one conserved domain and clustered into group I have been reported in peanut (*Arachis hypogaea* L.) (Zhao et al., 2020).

The intron-exon structure seems to be diverse, starting from 1 to 9 introns. Diversification in intron-exon structure may play an important role in the evolution of genes (Xu et al., 2012). In this study, we have identified different intron constitutions within gene groups. Group I has four introns on average, group IIa has 5, group IIb has 4, and all other groups have an average of three introns. This suggests that the latter groups originated from group I and earlier subgroups of group II. The same trend of intron presence in different groups of *WRKY* members in eggplant was reported (Yang et al., 2020). A total of six motifs were detected; among them, two motifs were considered to be *WRKY* domains. The conserved domain "WRKYGQK" was almost included in all of the members; however, some of the *WRKY* members showed a slight variation as "WRKYGKK" domain as suggested by Liu et al. (2021) and Qu et al. (2021).

The *cis*-acting elements are involved in plant development and flowering induction. There are several elements, such as auxin, gibberellins, and jasmonic acid, which affect flowering-related traits. Jasmonates, cytokinins, and auxin control second bud growth in chrysanthemum (Sun et al., 2021). Auxin controls plant growth and development (Singla et al., 2006). The petal is the most important factor influencing the chrysanthemum's ornamental value. Auxin signaling is associated with petal growth in chrysanthemums (Wang J. et al., 2017). AuR is found in most *ClWRKY* members and may play a role in plant development and flowering induction. GRE was abundant in *ClWRKY*, an important hormone associated with flowering induction in chrysanthemum (Dong et al., 2017). CmBBX24 regulates flowering in chrysanthemum in conjunction with gibberellin synthesis (Zhu et al., 2020). In addition, the key recognition site for the *WRKY* gene, the W-box containing the TGAC sequence, was observed in most *ClWRKY* members. The W-box contributes under abiotic stress conditions (Dhatterwal et al., 2019) by influencing the early senescence of rice flag leaves (Liu et al., 2016). The presence of flower-associated *cis*-acting elements in most genes indicates the potential of *ClWRKY* to play a role in flowering control in *C. lavandulifolium.*

DNA methylation has an impact on growth and development, with CG content, or CpG, being the most abundant source of DNA methylation (Dong-ru et al., 2019). Various methods such as methylation-sensitive amplification polymorphism and MSP have been applied to methylated genes (Herman et al., 1996; Coronel et al., 2018). In this study, we have identified ten genes with strong CpG content in the promoter region (**Table 1**), which may cause epigenetic changes in *C. lavandulifolium* (Chip and Dresselhaus, 2019). Three genes having methylated sites, i.e., *ClWRKY31, ClWRKY100,* and *ClWRKY129*, were identified on the basis of MSP results. These genes can regulate gene expression and may play a role in flowering-related traits, as they are expressed significantly in capitulum tissues at the flowering stage. *CmMET1* can affect methylation and regulate candidate genes associated with flowering in *C. morifolium* (Kang et al., 2021). DNA methylation changes in the promoter and gene body regions can influence gene expression and function of the protein

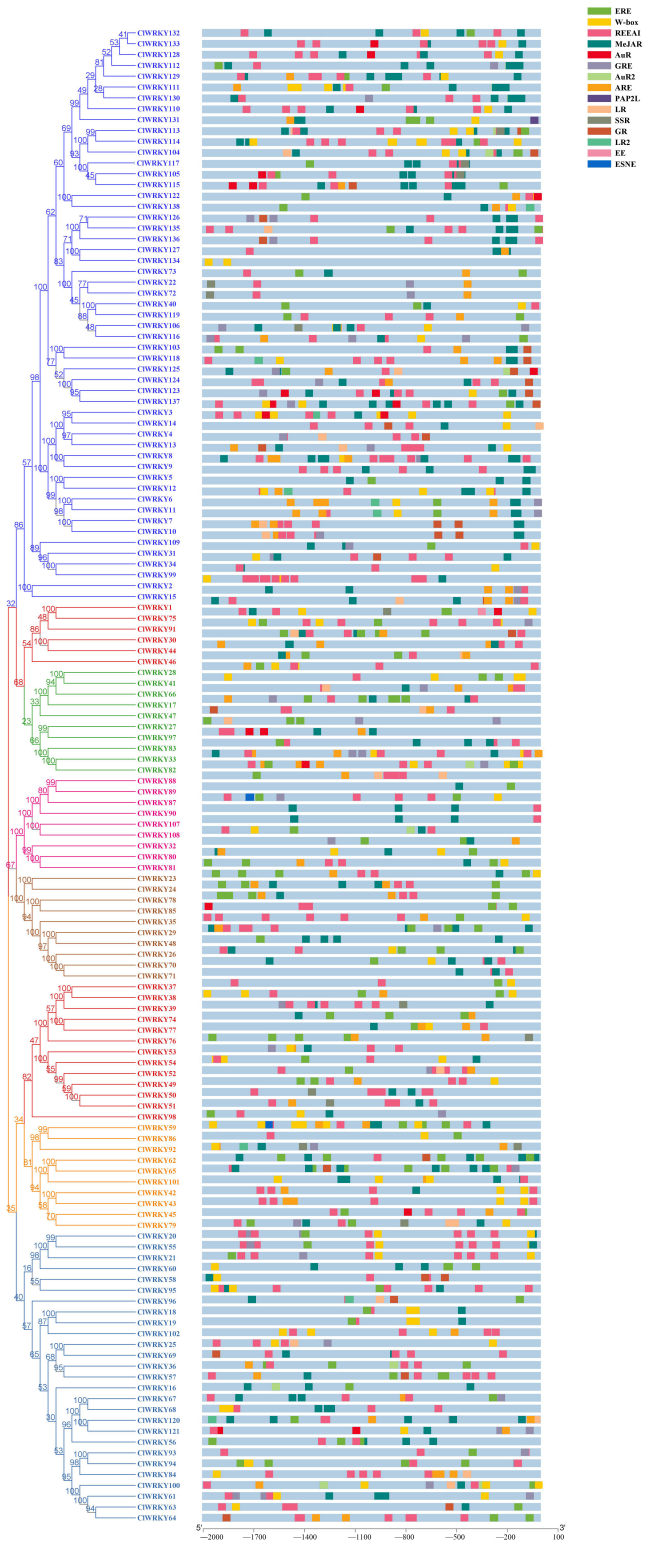


FIGURE 6 | *Cis*-acting element analysis in the promoter region of the *ClWRKY* family. The PlantCARE software was used to search the *cis*-acting elements. Abbreviation of important *cis*-acting elements are as follows: EE, endosperm expression; ARE, auxin-responsive elements; GR, gibberellin responsiveness; LR, light responsiveness; REEAI, regulatory element essential for anaerobic induction; AuR, elements involved in auxin responsiveness; SSR, seed-specific regulation; MeJAR, methyl jasmonate responsiveness; ERE, ethylene-responsive elements; GRE, gibberellin responsive elements; ESNE, endosperm specific negative expression; LR2, elements involved in light response; PAP2L, plant-AP2 such as AuR2, part of auxin responsiveness; and W-box, recognition site of *WRKY* TFs.

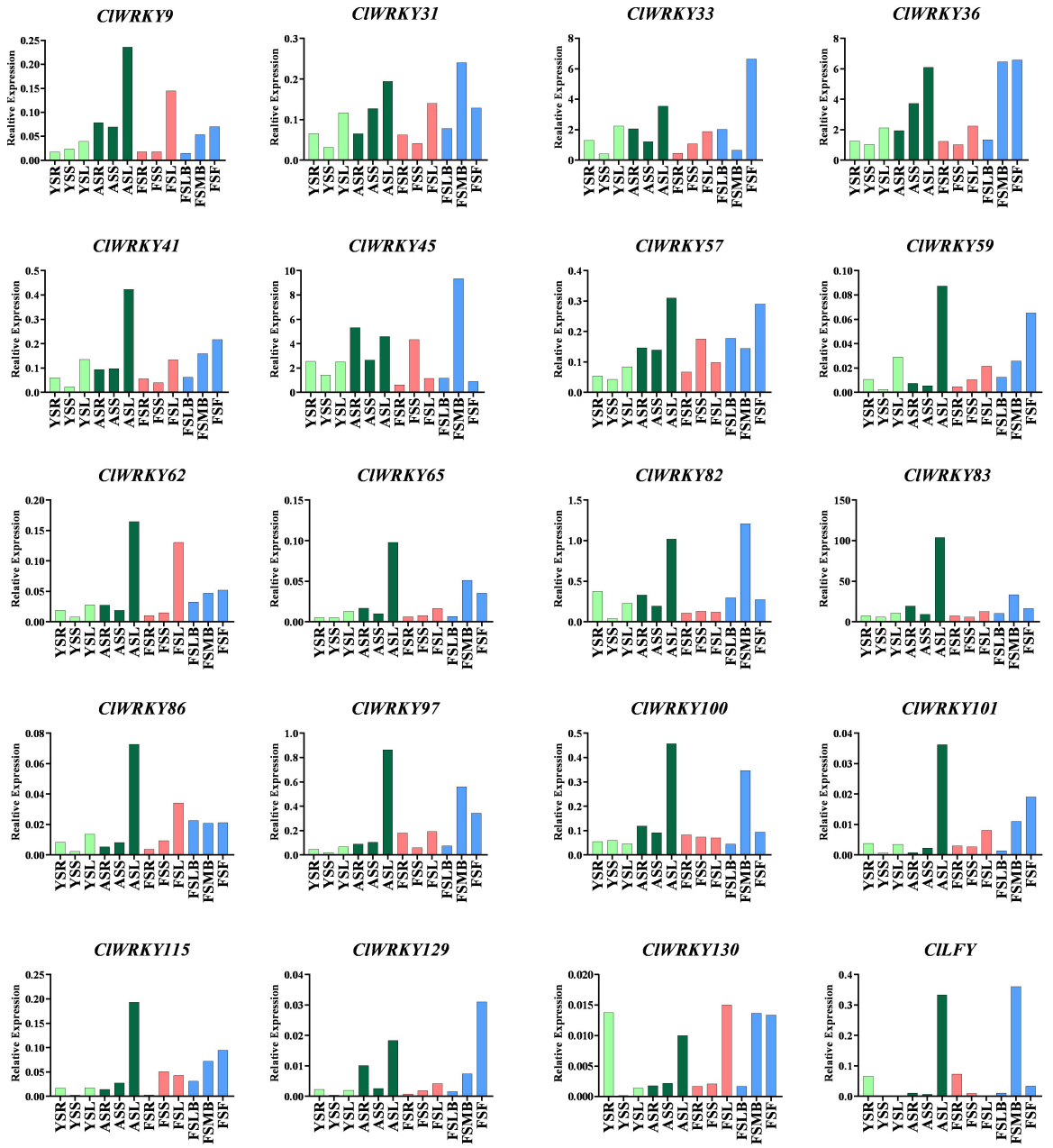


FIGURE 7 | Expression analysis of the selected *ClWRKY* genes in different tissues (capitulum, leaf, stem, and root) at different stages (young, adult, and flowering). Expression of 19 genes: 10 genes were selected due to the presence of CpG island, nine genes were selected based on their close association with important *Arabidopsis* genes found in the phylogenetic tree built between *C. lavandulifolium* and *A. thaliana*. Abbreviations of different growth stages and tissues are as follows: YSR, young stage root; YSS, young stage stem; YSL, young stage leaf; ASR, adult stage root; ASS, adult stage stem; ASL, adult stage leaf; FSR, flowering stage root; FSS, flowering stage stem; FSL, flowering stage leaf; FSLB, Flowering stage little buds; FSMB, Flowering stage buds; FSF, Flowering stage first flower.

(Miguel and Marum, 2011), the morphology of floral characters, photoperiod-related sensitivity, and fruit ripening (Sun et al., 2014; Xiao et al., 2020). We studied only four stages to detect methylation, and it will be better to explore the methylation status in more stages in the future.

This study focuses on the expression of *ClWRKY* members to find their role in plant development, especially their influence over flower induction. The expression analysis revealed variations of different genes in different organs. *ClWRKY45* showed significant expression at the flowering stage in the capitulum tissue, which may be related to some morphological characters and floral organ development. In the same fashion, *ClWRKY45* clustered with *Arabidopsis* At*WRKY71*, a prominent gene that controls flowering characteristics (Yu et al., 2016, 2018),

providing strong evidence for our results. Moreover, *ClWRKY36* and *ClWRKY57* are significantly expressed in capitulum at the flowering stages and have close homology with *Arabidopsis AtWRKY2* and *AtWRKY34,* indicating the role of *ClWRKY36* and *ClWRKY57* in flowering, as *Arabidopsis* homologs are the key regulators of pollen development (Lei et al., 2017). *ClWRKY59* is expressed significantly in the first flower tissue and can positively regulate flowering, as *ClWRKY59* homolog, *AtWRKY75*, promoted flowering in *Arabidopsis* (Zhang et al., 2018). *ClWRKY82* exhibited higher expression in the median bud stage, which indicates its potential to play a role in flowering regulation. *AtWRKY7,* a *ClWRKY82* homolog, regulates plant growth (Chen et al., 2019). The qRT-PCR results indicate that *ClWRKY83* is a stable gene as it expresses across all the tissues at different stages of plant life and may play a role in adaptability. The higher expression of three genes in capitulum tissues, including *ClWRKY31, ClWRKY100,* and *ClWRKY129*, which are methylated as well, can play an important role in flowering-related traits. These results suggest that *ClWRKY33, ClWRKY36,* and *ClWRKY45,* along with *ClWRKY31*, *ClWRKY100*, and *ClWRKY129*, can play a crucial role in the genetic and epigenetic improvement of the chrysanthemum. *ClWRKY45, ClWRKY36*, and *ClWRK129*, in particular, have the potential to alter flowering-related traits. The overall investigation of the *ClWRKY* domain is providing theoretical insight into *C. lavandulifolium* transcriptional activation and inactivation mechanisms for the improvement of plant development.

CONCLUSION

This study aimed to identify the *WRKY* gene family in *C. lavandulifolium* and explore the potential of *WRKY* genes for genetic improvement. A total of 138 *WRKY* genes were analyzed in *C. lavandulifolium*. All the data results are unique to explore the underlying mechanism of the WRKY TF in *C. lavandulifolium* plant improvement. The gene structure analysis suggested the vast variation among *WRKY* members. The *WRKY* domain remains conserved across all the members of this family. Ten *WRKY* genes possessed CG-rich content in the promoter region where three genes exhibit methylation status. This research can provide a basis to study the role of *WRKY* genes in the improvement of chrysanthemum plants, especially in flowering traits.

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

WZ designed the research project, coordinated, and supervised the study. MA performed all the experiments and wrote the manuscript. MA and KD did most of the experimental work and analyzed the data. WuY, AP, and WaY performed specific experiments during the research. All authors read and approved the final manuscript.

FUNDING

This research was supported by the National Key Research and Development Plan (2018YFD1000403).

ACKNOWLEDGMENTS

We thank all the colleagues who helped them a lot during this research. We specially thank WZ and KD for their support and timely response.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.861193/full#supplementary-material

REFERENCES

Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.

Bailey, T. L., and Grant, C. E. (2021). SEA: Simple Enrichment Analysis of motifs. *bioRxiv* [Preprint]. doi: 10.1101/2021.09.02.458722

Bailey, T. L., Johnson, J., Grant, C. E., and Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Res.* 43, W39–W49.

Bi, C., Xu, Y., Ye, Q., Yin, T., and Ye, N. (2016). Genome-wide identification and characterization of WRKY gene family in Salix suchowensis. *PeerJ* 4:e2437. doi: 10.7717/peerj.2437

Bi, M., Li, X., Yan, X., Liu, D., Gao, G., Zhu, P., et al. (2021). Chrysanthemum WRKY15-1 promotes resistance to Puccinia horiana Henn. *via* the salicylic acid signaling pathway. *Hortic Res.* 8:6. doi: 10.1038/s41438-020-00436-4

Brand, L. H., Fischer, N. M., Harter, K., Kohlbacher, O., and Wanke, D. (2013). Elucidating the evolutionary conserved DNA-binding specificities of WRKY transcription factors by molecular dynamics and *in vitro* binding assays. *Nucleic Acids Res.* 41, 9764–9778. doi: 10.1093/nar/gkt732

Cai, Y., Chen, X., Xie, K., Xing, Q., Wu, Y., Li, J., et al. (2014). Dlf1, a WRKY transcription factor, is involved in the control of flowering time and plant height in rice. *PLoS One* 9:e102529. doi: 10.1371/journal.pone.0102529

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, H., Lai, Z., Shi, J., Xiao, Y., Chen, Z., and Xu, X. (2010). Roles of arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biol.* 10:281. doi: 10.1186/1471-2229-10-281

Chen, W., Hao, W. J., Xu, Y. X., Zheng, C., Ni, D. J., Yao, M. Z., et al. (2019). Isolation and characterization of CsWRKY7, a subgroup IId WRKY transcription factor from Camellia sinensis, linked to development in arabidopsis. *Int. J. Mol. Sci.* 20:2815. doi: 10.3390/ijms20112815

Chip, I., and Dresselhaus, T. (2019). Chapter 12 and Reproductive Development in Cereals Using Chromatin. *Cereal Genomics* 2072, 141–156. doi: 10.1007/978-1-4939-9865-4_12

Coronel, C. J., González, A. I., Ruiz, M. L., and Polanco, C. (2018). Analysis of somaclonal variation in transgenic and regenerated plants of Arabidopsis thaliana using methylation related metAFLP and TMD markers. *Plant Cell Rep.* 37, 137–152. doi: 10.1007/s00299-017-2217-x

Dhatterwal, P., Basu, S., Mehrotra, S., and Mehrotra, R. (2019). Genome wide analysis of W-box element in Arabidopsis thaliana reveals TGAC motif with genes down regulated by heat and salinity. *Sci Rep.* 9:1681. doi: 10.1038/s41598-019-38757-7

Dong, B., Deng, Y., Wang, H., Gao, R., Stephen, G. K., Chen, S., et al. (2017). Gibberellic acid signaling is required to induce flowering of chrysanthemums grown under both short and long days. *Int. J. Mol. Sci.* 18:1259. doi: 10.3390/ijms18061259

Dong-ru, K., Si-lan, D., Kang, G., Fan, Z., and Hong, L. (2019). Morphological variation of Chrysanthemum lavandulifolium induced by 5-azaC treatment. *Sci. Hortic.* 257:108645. doi: 10.1016/j.scienta.2019.108645

Duan, M. R., Nan, J., Liang, Y. H., Mao, P., Lu, L., Li, L., et al. (2007). DNA binding mechanism revealed by high resolution crystal structure of Arabidopsis thaliana WRKY1 protein. *Nucleic Acids Res.* 35, 1145–1154. doi: 10.1093/nar/gkm001

Fan, Q., Song, A., Jiang, J., Zhang, T., Sun, H., Wang, Y., et al. (2016). CmWRKY1 enhances the dehydration tolerance of chrysanthemum through the regulation of ABA-aociated genes. *PLoS One* 11:e0150572. doi: 10.1371/journal.pone.0150572

Fan, Q., Song, A., Xin, J., Chen, S., Jiang, J., Wang, Y., et al. (2015). CmWRKY15 facilitates Alternaria tenuissima infection of chrysanthemum. *PLoS One* 10:e0143349. doi: 10.1371/journal.pone.0143349

Fu, J., Wang, Y., Huang, H., Zhang, C., and Dai, S. (2013). Reference gene selection for RT-qPCR analysis of Chrysanthemum lavandulifolium during its flowering stages. *Mol. Breed.* 31, 205–215. doi: 10.1007/s11032-012-9784-x

Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., and Noble, W. S. (2007). Quantifying similarity between motifs. *Genome Biol.* 8:R24, doi: 10.1186/gb-2007-8-2-r24

He, L., Wu, Y. H., Zhao, Q., Wang, B., Liu, Q. L., and Zhang, L. (2018). Chrysanthemum DgWRKY2 gene enhances tolerance to salt stress in transgenic chrysanthemum. *Int. J. Mol. Sci.* 19:2062. doi: 10.3390/ijms19072062

Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D., and Baylin, S. B. (1996). Methylation-specific PCR: A novel PCR assay for methylation status of CpG islands (DNA methylation/tumor suppressor genes/pl6/p15). *Med. Sci.* 93, 9821–9826. doi: 10.1073/pnas.93.18.9821

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Huang, H., Cao, H., Niu, Y., and Dai, S. (2012). Expression Analysis of Nudix Hydrolase Genes in Chrysanthemum lavandulifolium. *Plant Mol. Biol. Rep.* 30, 973–982. doi: 10.1007/s11105-011-0401-7

Ishiguro, S., and Nakamura, K. (1994). Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and β-amylase from sweet potato. *MGG Mol. Gen. Genet.* 244, 563–571. doi: 10.1007/BF00282746

Jaffar, M. A., Song, A., Faheem, M., Chen, S., Jiang, J., Liu, C., et al. (2016). Involvement of CmWRKY10 in drought tolerance of chrysanthemum through the ABA-signaling pathway. *Int. J. Mol. Sci.* 17:693. doi: 10.3390/ijms17050693

Kang, D. R., Zhu, Y., Li, S. L., Ai, P. H., Khan, M. A., Ding, H. X., et al. (2021). Transcriptome analysis of differentially expressed genes in chrysanthemum MET1 RNA interference lines. *Physiol. Mol. Biol. Plants* 27, 1455–1468. doi: 10.1007/s12298-021-01022-1

Kumar, K., Srivastava, V., Purayannur, S., Kaladhar, V. C., Cheruvu, P. J., and Verma, P. K. (2016). WRKY domain-encoding genes of a crop legume chickpea (Cicer arietinum): Comparative analysis with Medicago truncatula WRKY family and characterization of group-III gene(s). *DNA Res.* 23, 225–239. doi: 10.1093/dnares/dsw010

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

Lei, R., Li, X., Ma, Z., Lv, Y., Hu, Y., and Yu, D. (2017). Arabidopsis WRKY2 and WRKY34 transcription factors interact with VQ20 protein to modulate pollen development and function. *Plant J.* 91, 962–976. doi: 10.1111/tpj.13619

Lei, Y., Sun, Y., Wang, B., Yu, S., Dai, H., Li, H., et al. (2020). Woodland strawberry WRKY71 acts as a promoter of flowering *via* a transcriptional regulatory cascade. *Hortic Res.* 7:137. doi: 10.1038/s41438-020-00355-4

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, H. L., Guo, D., Yang, Z. P., Tang, X., and Peng, S. Q. (2014). Genome-wide identification and characterization of WRKY gene family in Hevea brasiliensis. *Genomics* 104, 14–23. doi: 10.1016/j.ygeno.2014.04.004

Li, P., Song, A., Gao, C., Jiang, J., Chen, S., Fang, W., et al. (2015a). The over-expression of a chrysanthemum WRKY transcription factor enhances aphid resistance. *Plant Physiol. Biochem.* 95, 26–34. doi: 10.1016/j.plaphy.2015.07.002

Li, P., Song, A., Gao, C., Wang, L., Wang, Y., Sun, J., et al. (2015b). Chrysanthemum WRKY gene CmWRKY17 negatively regulates salt stress tolerance in transgenic chrysanthemum and Arabidopsis plants. *Plant Cell Rep.* 34, 1365–1378. doi: 10.1007/s00299-015-1793-x

Li, S., Fu, Q., Chen, L., Huang, W., and Yu, D. (2011). Arabidopsis thaliana WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. *Planta* 233, 1237–1252. doi: 10.1007/s00425-011-1375-2

Li, S., Fu, Q., Huang, W., and Yu, D. (2009). Functional analysis of an Arabidopsis transcription factor WRKY25 in heat stress. *Plant Cell Rep.* 28, 683–693. doi: 10.1007/s00299-008-0666-y

Li, W., Wang, H., and Yu, D. (2016). Arabidopsis WRKY Transcription Factors WRKY12 and WRKY13 Oppositely Regulate Flowering under Short-Day Conditions. *Mol. Plant* 9, 1492–1503. doi: 10.1016/j.molp.2016.08.003

Liang, Q. Y., Wu, Y. H., Wang, K., Bai, Z. Y., Liu, Q. L., Pan, Y. Z., et al. (2017). Chrysanthemum WRKY gene DgWRKY5 enhances tolerance to salt stress in transgenic chrysanthemum. *Sci. Rep.* 7:4799. doi: 10.1038/s41598-017-05170-x

Lin, L. Z., and Harnly, J. M. (2010). Identification of the phenolic components of chrysanthemum flower (Chrysanthemum morifolium Ramat). *Food Chem.* 120, 319–326. doi: 10.1016/j.foodchem.2009.09.083

Liu, J., Wang, X., Chen, Y., Liu, Y., Wu, Y., Ren, S., et al. (2021). Identification, evolution and expression analysis of WRKY gene family in Eucommia ulmoides. *Genomics* 113, 3294–3309. doi: 10.1016/j.ygeno.2021.05.011

Liu, L., Xu, W., Hu, X., Liu, H., and Lin, Y. (2016). W-box and G-box elements play important roles in early senescence of rice flag leaf. *Sci. Rep.* 6:20881. doi: 10.1038/srep20881

Liu, Q. L., Xu, K. D., Pan, Y. Z., Jiang, B. B., Liu, G. L., Jia, Y., et al. (2014). Functional Analysis of a Novel Chrysanthemum WRKY Transcription Factor Gene Involved in Salt Tolerance. *Plant Mol. Biol. Rep.* 32, 282–289. doi: 10.1007/s11105-013-0639-3

Liu, Q. L., Zhong, M., Li, S., Pan, Y. Z., Jiang, B. B., Jia, Y., et al. (2013). Overexpression of a chrysanthemum transcription factor gene, DgWRKY3, intobacco enhances tolerance to salt stress. *Plant Physiol. Biochem.* 69, 27–33. doi: 10.1016/j.plaphy.2013.04.016

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.2001.1262

Luo, X., Sun, X., Liu, B., Zhu, D., Bai, X., Cai, H., et al. (2013). Ectopic Expression of a WRKY Homolog from Glycine soja Alters Flowering Time in Arabidopsis. *PLoS One* 8:e73295. doi: 10.1371/journal.pone.0073295

Ma, Z., Li, W., Wang, H., and Yu, D. (2020). WRKY transcription factors WRKY12 and WRKY13 interact with SPL10 to modulate age-mediated flowering. *J. Integr. Plant Biol.* 62, 1659–1673. doi: 10.1111/jipb.12946

Miguel, C., and Marum, L. (2011). An epigenetic view of plant cells cultured *in vitro*: Somaclonal variation and beyond. *J. Exp. Bot.* 62, 3713–3725. doi: 10.1093/jxb/err155

Nan, H., and Gao, L. Z. (2019). Genome-wide analysis of WRKY genes and their response to hormone and mechanic stresses in carrot. *Front. Genet.* 10:363. doi: 10.3389/fgene.2019.00363

Nguyen, T. K., and Lim, J. H. (2019). Tools for Chrysanthemum genetic research and breeding: Is genotyping-by-sequencing (GBS) the best approach? *Hortic. Environ. Biotechnol.* 60, 625–635. doi: 10.1007/s13580-019-00160-6

Phukan, U. J., Jeena, G. S., and Shukla, R. K. (2016). WRKY transcription factors: Molecular regulation and stress responses in plants. *Front. Plant Sci.* 7:760. doi: 10.3389/fpls.2016.00760

Qu, R., Cao, Y., Tang, X., Sun, L., Wei, L., and Wang, K. (2021). Identification and expression analysis of the WRKY gene family in Isatis indigotica. *Gene* 783:145561. doi: 10.1016/j.gene.2021.145561

Ramos, R. N., Martin, G. B., Pombo, M. A., and Rosli, H. G. (2021). WRKY22 and WRKY25 transcription factors are positive regulators of defense responses in Nicotiana benthamiana. *Plant Mol Biol.* 105, 65–82. doi: 10.1007/s11103-020-01069-w

Rinerson, C. I., Rabara, R. C., Tripathi, P., Shen, Q. J., and Rushton, P. J. (2015). The evolution of WRKY transcription factors. *BMC Plant Biol.* 15:66. doi: 10.1186/s12870-015-0456-y

Rombauts, S., Déhais, P., Van Montagu, M., and Rouzé, P. (1999). PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* 27, 295–296. doi: 10.1093/nar/27.1.295

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

Scarpeci, T. E., Zanor, M. I., Mueller-Roeber, B., and Valle, E. M. (2013). Overexpression of AtWRKY30 enhances abiotic stress tolerance during early growth stages in Arabidopsis thaliana. *Plant Mol Biol.* 83, 265–277. doi: 10.1007/s11103-013-0090-8

Shinoyama, H., Aida, R., Ichikawa, H., Nomura, Y., and Mochizuki, A. (2012). Genetic engineering of chrysanthemum (Chrysanthemum morifolium): Current progress and perspectives. *Plant Biotechnol.* 29, 323–337. doi: 10.5511/plantbiotechnology.12.0521a

Singla, B., Chugh, A., Khurana, J. P., and Khurana, P. (2006). An early auxin-responsive Aux/IAA gene from wheat (Triticum aestivum) is induced by epibrassinolide and differentially regulated by light and calcium. *J. Exp. Bot.* 57, 4059–4070. doi: 10.1093/jxb/erl182

Soares, J. M., Weber, K. C., Qiu, W., Stanton, D., Mahmoud, L. M., Wu, H., et al. (2020). The vascular targeted citrus FLOWERING LOCUS T3 gene promotes non-inductive early flowering in transgenic Carrizo rootstocks and grafted juvenile scions. *Sci. Rep.* 10:21404. doi: 10.1038/s41598-020-78417-9

Song, A., Li, P., Jiang, J., Chen, S., Li, H., Zeng, J., et al. (2014). Phylogenetic and transcription analysis of chrysanthemum WRKY transcription factors. *Int. J. Mol. Sci.* 15, 14442–14455. doi: 10.3390/ijms150814442

Song, H., Wang, P., Hou, L., Zhao, S., Zhao, C., Xia, H., et al. (2016a). Global analysis of WRKY genes and their response to dehydration and salt stress in soybean. *Front. Plant Sci.* 7:9. doi: 10.3389/fpls.2016.00009

Song, H., Wang, P., Lin, J. Y., Zhao, C., Bi, Y., and Wang, X. (2016b). Genome-wide identification and characterization of WRKY gene family in peanut. *Front. Plant Sci.* 7:534. doi: 10.3389/fpls.2016.00534

Song, H., Wang, P., Nan, Z., and Wang, X. (2014). The WRKY transcription factor genes in lotus japonicus. *Int. J. Genom.* 2014:420128. doi: 10.1155/2014/420128

Song, Y., Cui, H., Shi, Y., Xue, J., Ji, C., Zhang, C., et al. (2020). Genome-wide identification and functional characterization of the Camelina sativa WRKY gene family in response to abiotic stress. *BMC Genom.* 21:786. doi: 10.1186/s12864-020-07189-3

Sun, D., Zhang, L., Yu, Q., Zhang, J., Li, P., Zhang, Y., et al. (2021). Integrated Signals of Jasmonates, Sugars, Cytokinins and Auxin Influence the Initial Growth of the Second Buds of Chrysanthemum after Decapitation. *Biology* 10:440 doi: 10.3390/biology10050440

Sun, H., Guo, Z., Gao, L., Zhao, G., Zhang, W., Zhou, R., et al. (2014). DNA methylation pattern of Photoperiod-B1 is associated with photoperiod insensitivity in wheat (Triticum aestivum). *New Phytol.* 204, 682–692. doi: 10.1111/nph.12948

Turck, F., Zhou, A., and Somssich, I. E. (2004). Stimulus-dependent, promoter-specific binding of transcription factor WRKY1 to its native promoter and the defense-related gene PcPR1-1 in parsley. *Plant Cell* 16, 2573–2585. doi: 10.1105/tpc.104.024810

Wang, J., Wang, H., Ding, L., Song, A., Shen, F., Jiang, J., et al. (2017). Transcriptomic and hormone analyses reveal mechanisms underlying petal elongation in Chrysanthemum morifolium 'Jinba.' *Plant Mol. Biol.* 93, 593–606. doi: 10.1007/s11103-017-0584-x

Wang, K., Wu, Y. H., Tian, X. Q., Bai, Z. Y., Liang, Q. Y., Liu, Q. L., et al. (2017). Overexpression of DgWRKY4 enhances salt tolerance in chrysanthemum seedlings. *Front. Plant Sci.* 8:1592. doi: 10.3389/fpls.2017.01592

Wang, Y., Li, Y., He, S. P., Gao, Y., Wang, N. N., Lu, R., et al. (2019). A cotton (Gossypium hirsutum) WRKY transcription factor (GhWRKY22) participates in regulating anther/pollen development. *Plant Physiol. Biochem.* 141, 231–239. doi: 10.1016/j.plaphy.2019.06.005

Warmerdam, S., Sterken, M. G., Sukarta, O. C. A., Van Schaik, C. C., Oortwijn, M. E. P., Lozano-Torres, J. L., et al. (2020). The TIR-NB-LRR pair DSC1 and WRKY19 contributes to basal immunity of Arabidopsis to the root-knot nematode Meloidogyne incognita. *BMC Plant Biol.* 20:73. doi: 10.1186/s12870-020-2285-x

Wen, X., Li, J., Wang, L., Lu, C., Gao, Q., Xu, P., et al. (2022). The chrysanthemum lavandulifolium genome and the molecular mechanism underlying diverse capitulum types. *Hortic Res.* 9:uhab022. doi: 10.1093/hr/uhab022

Wu, J., Zhang, Y., Zhang, H., Huang, H., Folta, K. M., and Lu, J. (2010). Whole genome wide expression profiles of Vitis amurensis grape responding to downy mildew by using Solexa sequencing technology. *BMC Plant Biol.* 10:234. doi: 10.1186/1471-2229-10-234

Xiao, K., Chen, J., He, Q., Wang, Y., Shen, H., and Sun, L. (2020). DNA methylation is involved in the regulation of pepper fruit ripening and interacts with phytohormones. *J. Exp. Bot.* 71, 1928–1942. doi: 10.1093/jxb/eraa003

Xu, G., Guo, C., Shan, H., and Kong, H. (2012). Divergence of duplicate genes in exon-intron structure. *Proc. Natl. Acad. Sci. U.S.A.* 109, 1187–1192. doi: 10.1073/pnas.1109047109

Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., et al. (2005). Solution structure of an Arabidopsis WRKY DNA binding domain. *Plant Cell* 17, 944–956. doi: 10.1105/tpc.104.026435

Yang, C., Ye, Y., Song, C., Chen, D., Jiang, B., and Wang, Y. (2016). Cloning and functional identification of the AcLFY gene in Allium cepa. *Biochem. Biophys. Res. Commun.* 473, 1100–1105. doi: 10.1016/j.bbrc.2016.04.022

Yang, L., Fu, J., Qi, S., Hong, Y., Huang, H., and Dai, S. (2017). Molecular cloning and function analysis of ClCRY1a and ClCRY1b, two genes in Chrysanthemum lavandulifolium that play vital roles in promoting floral transition. *Gene* 617, 32–43. doi: 10.1016/j.gene.2017.02.020

Yang, Y., Liu, J., Zhou, X., Liu, S., and Zhuang, Y. (2020). Identification of WRKY gene family and characterization of cold stress-responsive WRKY genes in eggplant. *PeerJ* 8:e8777. doi: 10.7717/peerj.8777

Yokotani, N., Shikata, M., Ichikawa, H., Mitsuda, N., Ohme-Takagi, M., Minami, E., et al. (2018). OsWRKY24, a blast-disease responsive transcription factor, positively regulates rice disease resistance. *J. Gen. Plant Pathol.* 84, 85–91. doi: 10.1007/s10327-018-0768-5

Yu, Y., Liu, Z., Wang, L., Kim, S. G., Seo, P. J., Qiao, M., et al. (2016). WRKY71 accelerates flowering *via* the direct activation of FLOWERING LOCUS T and LEAFY in Arabidopsis thaliana. *Plant J.* 85, 96–106. doi: 10.1111/tpj.13092

Yu, Y., Wang, L., Chen, J., Liu, Z., Park, C. M., and Xiang, F. (2018). WRKY71 Acts Antagonistically Against Salt-Delayed Flowering in Arabidopsis thaliana. *Plant Cell Physiol.* 59, 414–422. doi: 10.1093/pcp/pcx201

Yuan, H., Guo, W., Zhao, L., Yu, Y., Chen, S., Tao, L., et al. (2021). Genome-wide identification and expression analysis of the WRKY transcription factor family in flax (Linum usitatissimum L.). *BMC Genom.* 22:375. doi: 10.1186/s12864-021-07697-w

Yue, J., Zhu, C., Zhou, Y., Niu, X., Miao, M., Tang, X., et al. (2018). Transcriptome analysis of differentially expressed unigenes involved in flavonoid biosynthesis during flower development of Chrysanthemum morifolium 'Chuju.' *Sci. Rep.* 8:13414 . doi: 10.1038/s41598-018-31831-6

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, W., Gao, T., Li, P., Tian, C., Song, A., Jiang, J., et al. (2020). Chrysanthemum CmWRKY53 negatively regulates the resistance of chrysanthemum to the aphid Macrosiphoniella sanborni. *Hortic Res.* 7:109, doi: 10.1038/s41438-020-0334-0

Zhao, N., He, M., Li, L., Cui, S., Hou, M., Wang, L., et al. (2020). Identification and expression analysis of WRKY gene family under drought stress in peanut (Arachis hypogaea L.). *PLoS One* 15:e0231396. doi: 10.1371/journal.pone.0231396

Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., et al. (2008). Soybean WRKY-type transcription factor genes, GmWRKY13, GmWRKY21, and GmWRKY54, confer differential tolerance to abiotic stresses in transgenic Arabidopsis plants. *Plant Biotechnol. J.* 6, 486–503. doi: 10.1111/j.1467-7652.2008.00336.x

Zhou, X., Jiang, Y., and Yu, D. (2011). WRKY22 transcription factor mediates dark-induced leaf senescence in Arabidopsis. *Mol. Cells* 31, 303–313. doi: 10.1007/s10059-011-0047-1

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: 30 June 2022
doi: 10.3389/fpls.2022.893245

Transcriptome Analysis of Short-Day Photoperiod Inducement in Adzuki Bean (*Vigna angularis* L.) Based on RNA-Seq

Weixin Dong[1,2†], *Dongxiao Li*[1†], *Lei Zhang*[1†], *Baozhong Yin*[1] *and Yuechen Zhang*[1*]

[1] *State Key Laboratory of North China Crop Improvement and Regulation/Key Laboratory of Crop Growth Regulation of Hebei Province/College of Agronomy, Hebei Agricultural University, Baoding, China,* [2] *Hebei Open University, Shijiazhuang, China*

The flowering characteristics of adzuki bean are influenced by several environmental factors. Light is an important ecological factor that induces flowering in adzuki bean, but to date, there have been few reports on the transcriptomic features of photoperiodic regulation of adzuki bean flowering. This study is based on RNA sequencing (RNA-seq) techniques to elucidate the expression of light-related regulatory genes under short-day photoperiod inducement of adzuki bean flowering, providing an important theoretical basis for its accelerated breeding. Short-day photoperiod inducement of 10 h was conducted for 5 day, 10 day, and 15 day periods on "Tang shan hong xiao dou" varieties, which are more sensitive to short-day photoperiod conditions than the other varieties. Plants grown under natural light (14.5 h) for 5 days, 10 days, and 15 days were used as controls to compare the progress of flower bud differentiation and flowering characteristics. The topmost unfolded functional leaves were selected for transcriptome sequencing and bioinformatics analysis. The short-day photoperiod inducement promoted flower bud differentiation and advanced flowering time in adzuki bean. Transcriptomic analysis revealed 5,608 differentially expressed genes (DEGs) for the combination of CK-5d vs. SD-5d, CK-10d vs. SD-10d, and CK-15d vs. SD-15d. The three groups of the DEGs were analyzed using the Gene Ontology (GO) and the Kyoto Encyclopedia of Genomes and Genomes (KEGG) databases; the DEGs were associated with flowering, photosystem, and the circadian rhythm and were mainly concentrated in the hormone signaling and metabolism, circadian rhythm, and antenna protein pathways; So, 13 light-related genes across the three pathways were screened for differential and expression characteristics. Through the functional annotations of orthologs, these genes were related to flowering, which were supposed to be good candidate genes in adzuki bean. The findings provide a deep understanding of the molecular mechanisms of adzuki bean flowering in response to short-day photoperiod inducement, which laid a foundation for the functional verification of genes in the next step, and provide an important reference for the molecular breeding of adzuki bean.

Keywords: adzuki bean, short-day photoperiod inducement, high-throughput sequencing, transcriptomic analysis, qRT-PCR

OPEN ACCESS

Edited by:
Leo Marcelis, Wageningen University and Research, Netherlands

Reviewed by:
Rong Zhou, Aarhus University, Denmark
Geoffrey Thomson, Yale University, United States

***Correspondence:**
Yuechen Zhang
dongweixin.yuxin@163.com; Zhangyc1964@126.com

†These authors have contributed equally to this work

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

Received: *10 March 2022*
Accepted: *02 June 2022*
Published: *30 June 2022*

Citation:
Dong W, Li D, Zhang L, Yin B and Zhang Y (2022) Transcriptome Analysis of Short-Day Photoperiod Inducement in Adzuki Bean (Vigna angularis L.) Based on RNA-Seq. Front. Plant Sci. 13:893245. doi: 10.3389/fpls.2022.893245

INTRODUCTION

Adzuki bean is a grain crop in China with high nutritional and medicinal values and plays an important role in the dietary structure of residents. Its annual export volume is between 50,000 and 80,000 tons (Yin et al., 2019). However, the production of adzuki bean across the country has been limited owing to severe floral bud abscission, low and unstable yield, and restricted cultivation area; these are related to environmental factors during the floral development of adzuki bean (Liu et al., 2015). Light is an important ecological factor that induces flowering in plants, and the flowering time and the flower developmental stages are regulated by the duration of sunlight. This process is closely related to the regulation of flower development, structural changes and related physiological processes, and key genes. A deeper understanding of the molecular mechanisms underlying the photoperiodic flower developmental process could help systematically elucidate the biological mechanisms of photoperiodism in adzuki bean (Song Y. H. et al., 2013; Shrestha et al., 2014).

Plant flowering is jointly regulated by photoperiodic, autonomous, vernalization, and gibberellin pathways. Among them, the photoperiodic pathway is the most important (Pajoro et al., 2014; Kong et al., 2016). The metabolic pathways in photoperiodic flower development and flowering have been extensively studied in different plants. Moreover, the molecular mechanisms underlying the regulation of photoperiodic flowering in *Arabidopsis* have been largely elucidated (Jeong and Clark, 2005; Izawa, 2007). FT protein as florigen can be transported from the leaf to the apical meristem of *Arabidopsis* (Corbesier et al., 2007; Cheng et al., 2018). In the apical meristem, FT protein interacts with FLOWERING LOCUS D (FD), 14-3-3 protein, etc., to form a complex that can promote the expression of downstream flowering genes (Taoka et al., 2011). In soybean, 10 E gene loci *(E1-E10)* and a *J* locus related to flowering time have been identified to date, of which *E7* and *E8* loci are involved in inhibiting flowering, but no related genes have been reported (Cober and Voldeng, 2001; Cober et al., 2010; Lin et al., 2021). The other E genes have been studied deeply, e.g., *E1* affects flowering time significantly, inhibits under short-day condition, and shows a bimodal diurnal variation pattern under the long-day condition, indicating its response to photoperiod and its dominant effect induced by long-day length (Xia Z. J. et al., 2012). Furthermore, E1 has been shown to act upstream of four FT-like genes in soybean (*Glyma.16G150700*, *Glyma.16G044100*, *Glyma.08G363100*, and *Glyma.18G298900*) and play an important role in promoting soybean flowering as members of the photoperiod gene family (Xia Z. J. et al., 2012; Zhai et al., 2014; Liu et al., 2018). *E2*, a homologous gene of *GI* and *GmGIa*, and *e2*/*e2* genotype induce early flowering of soybean by inducing the expression of *GmFT2a*, a flowering gene homolog (Watanabe et al., 2011). *E3* and *E4* are related to phytochrome A, which is associated with the regulation of flowering in soybean. Further analysis by mapping showed that *GmPhyA3* is a candidate gene at the *E3* locus and *GmphyA2* is encoded by E4 (Watanabe et al., 2009; Liu et al., 2008). *E5*, as a gene that does not work alone, may be generated by accidental outcrossing of pollens that have the E2 allele (Dissanayaka et al., 2016). *E6* is a new allele of the homolog *J* of *ELF3*, the early flowering gene of *Arabidopsis*. The long-juvenile (*LJ*), characterized by delayed flowering and maturation, could increase yield under short-day conditions and expand soybean cultivation in lower latitudes (Fang et al., 2021). In the soybean mature gene E9 locus, the abundance of FT2a is directly related to the change in the soybean flowering time (Zhao et al., 2016). At the end of the Gm08 chromosome, there is a new maturation site E10, where FT4 is the most likely candidate gene by protein interaction and plays a regulatory role in soybean flowering time (Samanfar et al., 2017). *J* is a homolog of *EARLY FLOWERING3* (*ELF3*) in *Arabidopsis thaliana*. The J protein binds to the legume-specific FLOWERING suppressor *E1* and downregulates the transcription of the *E1* promoter to release the T(FT) gene, which is the important flowering site and promotes flowering in a short time (Lu et al., 2017). In addition to the *E* genes, *FT* genes also play important roles in soybean flowering (Banfield et al., 1998; Su et al., 2021); two cognates of FT, *GmFT2a* and *GmFT5a*, upregulated under a short-day photoperiod has the same function as FT in *Arabidopsis thaliana*, but the effect of *GmFT5a* is more significant. High expression of *GmFT2a* and *GmFT5a* has been shown in the double mutations of two phytochrome A (*PHYA*) genes under long-day conditions with flowering times slightly earlier than those of the wild type under short-day photoperiods. This indicated that the expression of *GmFT2a* and *GmFT5a* under long daylength conditions is affected by the photoperiod regulation system mediated by *PHYA*. *GmFT2a* and *GmFT5a* coordinately control flowering and enable soybean to adapt to a wide range of photoperiodic environments (Kong et al., 2010). Similarly, in *C. cajan*, among 13 PEBP (FT) family genes, *CcFT6* and *CcFT8* act as crucial florigen T genes involving flowering regulation. The expression of *CcFT6* is upregulated under short-day conditions, whereas the *CcFT8* transcripts are enhanced under long-day photoperiods to promote the plant flowering process, suggesting that they regulate plant flowering through photoperiod-dependent pathways (Tribhuvan et al., 2020). Transcriptome analysis indicated that Mtvrn2 (VERNALISATION2-LIKEVEFS-box gene) mutants of legume *Medicago truncatula* precociously express FTa1 and other suites of genes including floral homeotic genes. Functional FTa1, an FT-like gene, work as the floral activator and has an elevated expression to regulate early flowering under long-day conditions (Jaudal et al., 2016). In chickpea, SNP loci from the CaCLV3_01 gene (*CLAVATA*) within a major CaqDTF1.1/CaqFN1.1 QTL has been identified and verified to be associated with flowering time (DTF) and flower number (FN) traits (Basu et al., 2019).

In summary, research studies on the screening and the functional analysis of flowering-related key genes in *Arabidopsis*, soybean, *C. cajan*, *Medicago*, chickpea, etc. have provided insights into photoperiodic responses and flowering mechanisms in leguminous plants. For adzuki bean, most studies on flower development have focused on morphology and phenotypic physiology. Previously, we found that the growth index of adzuki bean is largely determined by different short-day induction. The extended short-day induction duration decreased the chlorophyll, soluble sugar, and protein contents;

increased enzyme activity; and altered hormone dynamic balance homeostasis (Dong et al., 2016). Our microscopic observation also found that the differentiation phase of inflorescence primordia in flower buds is promoted in adzuki bean plants after short-day induction (Zhang et al., 2021). According to this, "Tang shan hong xiao dou", a late-maturing variety sensitive to short-day photoperiod conditions, was used to research functional leaf responses to light and flowering under different short-day induction conditions based on transcriptomic analysis (RNA-Seq). Taking flowering-related genes of soybean as a reference, we attempted to search for flowering responsive genes in the genome annotations of adzuki bean to deepen the understanding of the photoperiodic response and the molecular mechanism of flower development in adzuki bean. This study could help to further understand the flowering mechanism of this plant and provide an important theoretical basis for genetic improvement in the high light efficiency of adzuki bean.

MATERIALS AND METHODS

Experimental Material

"Tang shan hong xiao dou," a late maturing and photoperiod sensitive variety, was used as the experimental material, which was provided by the Research Institute of Grain and Oil Crops, Academy of Agricultural and Forestry Sciences of Hebei, China. The flowering time and maturity time are earlier of this variety than other regularly cultivated cultivars under short-day photoperiod inducement. The experiment was conducted at the teaching experimental station of Hebei Agricultural University, China (38°38, N, 115°E), during the growth seasons in 2021.

Short-Day Treatment of Plant Materials

This field experiment was performed at the teaching test base of Hebei Agricultural University. Each plot was designed to be 5 m in length, 1 m in width, and 20 cm in depth. Prior to seed sowing, 400 g compound fertilizer (N-P_2O_5-K_2O=15-15-15) was applied to each plot (5 ×1 m) and then plowed. The topsoil (0-20 cm) had 1.54% organic matter, 0.0987% total nitrogen, 84.97 ppm available N, 65.658 ppm available P, and 184.6 ppm available K. The seeds were sown on 24 June and planted in two rows in each plot with a spacing of 15 ×40 cm. When true leaves flatten under natural light (14.5 h), short-day treatment was started with stainless steel shelves placed on the upper side of the plot together with a covering opaque cloth for shading and fixing additional vents for adjusting temperature, CO_2 concentration, and relative humidity (**Table 1**). Then, short-day photoperiod was set to 10-h light/14-h dark. At the first leaf expansion stage, three photoperiod treatments were set up by sheltering the seedlings with an opaque cloth during the daytime. These treatments included natural photoperiod (14.5 h) with 5 days (control check-5d), 10 days (control check-10d), and 15 days (control check-15d), and short-day photoperiod treatments with 5 days (short day-5d), 10 days (short day-10d), and 15 days (short day-15d). After shading, all plants grew to the maturity stage under natural light with spraying 0.4% potassium dihydrogen phosphate. Irrigations were conducted at the first flowering stage and the pod setting stage. The diseases and insect pests were controlled according to the actual occurrence situation. The experimental treatment consisted of 18 plots with random block arrangement. Photoperiod treatment was initiated at 8:00 am and terminated at 6:00 pm with a 10-hour photoperiod duration. Intermediate leaves with the uppermost unfolded three complex leaves were taken at 9:00 am to be mixed from five to six plants after shading for 5 days, 10 days, and 15 days. Each treatment was repeated three times, and there were 18 samples marked as SD-5d-1, SD-5d-2, SD-5d-3, CK-5d-1, CK-5d-2, CK-5d-3; SD-10d-1, SD-10d-2, SD-10d-3, CK-10d-1, CK-10d-2, CK-10d-3; SD-15d-1, SD-15d-2, SD-15d-3, CK-15d-1, CK-15d-2, and CK-15d-3. After that, the samples were quick-frozen in liquid nitrogen and stored at −80°C. One week later, RNAs from the samples were extracted and subjected to sequencing analysis.

Measurement of Meteorological Factors

During the growth stage, light intensity at 20–30 cm above the canopy of the adzuki bean community in the plot was measured using a TES1332 illuminometer (TES1332, Taiwan, China). An LI-6400 portable photosynthetic apparatus (Li-COR, Lincoln, NE, USA) was used to measure the CO_2 concentration. A HOBO Pro V2 series instrument (U23-002, USA) was used to measure temperature and humidity, which were recorded automatically every hour. The data were exported after the testing was finished.

Measurement of Plant Growth Traits

Totally, nine plants with uniform growth in each plot were subjected to measure the plant growth traits after photoperiod treatments. The growth traits assessed included plant height, stem diameter, pinch number, leaf area, and shoot dry weight per plant. The plant height was obtained by measuring the distance (cm) from the cotyledon node to the growing point. The stem diameter of plants was determined by using an electronic vernier caliper (United Precision Machine Precision Measurement Limited Company, Shenzhen, China).

The number of main stem nodes was referred to as the number of first nodes starting from the true leaf. The leaf area was measured using a YMJ-B leaf area meter (from Hangzhou Hui Equipment limited company). The shoot dry weight per plant was measured using the conventional oven drying method.

Anatomical Observation and Slicing Production of Adzuki Bean Flower Buds

According to the division method of Jin and Wang (1995), flower bud differentiation of adzuki bean was divided into the pre-differentiation stage (PD), the inflorescence primordium differentiation stage (IP), the flower primordium differentiation stage (FP), the sepal primordium differentiation stage (SP), the petal primordium differentiation stage (PP), the stamen and carpel primordium differentiation stage (SCP), and the stamen and carpel structural differentiation stage (SCS). After shading was completed (about 20 days after seedling emergence), the top flower buds were taken for each treatment with 20 flower buds. After making paraffin sections, they were observed under a microscope (BS500/BS500-TR biological microscope). More than 70% of flower buds differentiated into a certain period was judged it has entered this developmental period.

TABLE 1 | Changes in field microclimate with shading treatment.

	Illumination intensity (lux)	CO_2 concentration(ppm)	Relative humidity (%)	Temperature(°C)
Atmosphere	57,963.33 ± 1,678.58a	451.53 ± 18.72a	78.82 ± 4.77a	26.59 ± 0.89a
Shading	10.63 ± 3.50b	463.88 ± 20.45a	82.33 ± 1.79a	27.28 ± 1.09a

Values are means ± S.E. The different small letters in the same column indicate statistical significance at 0.05 level by DMRT.

Production methods of paraffin sections include sampling, fixation, dehydration and transparent, paraffin-embedded, slicing and patch, and HE staining.

Investigation and Statistics of Flowering Characteristics

The time period when 50% of plants began to blossom was called the flowering period, when three representative plants were selected per treatment with three replicates and nine plants in total. Early flowering time was recorded, and the flowering promotion rate was calculated. The determination formula is as follows:

Flowering promotion rate (%) = [(Control the seedling emergence period to the flowering period-Different treatment of seedling period to flowering period time) × 100]/Control the seedling emergence period to the flowering period.

Early flowering time = Control the seedling emergence period to the flowering period – Different treatment of seedling period to flowering period time.

RNA Extraction and Library Construction

Total RNA was extracted using a mirVana miRNA Isolation Kit (Ambion) following the manufacturer's protocol. RNA integrity was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The samples with an RNA integrity number (RIN) $\geq$ 7 were subjected to the subsequent analysis. The libraries were constructed using a TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, USA) in accordance with the manufacturer's instructions. Then, these libraries were sequenced on the Illumina sequencing platform (HiSeqTM 2500 or Illumina HiSeq X Ten), and 125-bp/150-bp paired-end reads were generated.

Quality Control and Mapping

Raw data (raw reads) were processed using Trimmomatic (Bolger et al., 2014). The reads containing ploy-N and the low-quality reads were removed to obtain the clean reads. Then, the clean reads were mapped to the reference genome using hisat2 (Kim et al., 2015).

Analysis of Differentially Expressed Transcripts (DEGs)

Analysis of DEGs was performed through transcript-level quantification, and FPKM (Roberts et al., 2011) and read count value of each transcript (protein_coding) were calculated using bowtie2 (Langmead and Salzberg, 2012) and eXpress (Roberts and Pachter, 2013). DEGs (read counts >2) were identified using the DESeq (Anders and Huber, 2012) with functions estimate size factors and nbinomtest. $p < 0.05$ and a fold change >2 or fold change < 0.5 were set as the threshold for significantly differential expression. Differentially expressed genes (DEGs) were selected by comparing all genes expressed in adzuki bean leaves under different short-day photoperiod inducement conditions.

Functional Annotation Analysis of Differentially Expressed Genes

Differential transcripts with a $p < 0.05$ and a difference fold >2 were selected, GO annotation information using Blast2GO software for all the differential genes were obtained, and GO classification statistics for all the differential genes were used for GO using WEGO software. The GO annotation includes three main branches, namely, biological process, molecular function, and cellular component. Gene sequences were annotated to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database by BLAST software alignment, and then the organism's metabolic network was studied.

Quantitative Real-Time PCR (qRT-PCR) Analysis

To verify the accuracy of the adzuki bean sequencing results, 13 differentially expressed genes were verified again by qRT-PCR using new leaf samples under the same treatment. The specific steps were as follows: Total RNA was extracted from adzuki bean tissues with trizol reagent. Residual genomic DNA in the RNA samples was removed by RNase-free DNase following the manufacturer's protocol (New England Biolabs). In total, one microgram of DNase-treated total RNA was used to synthesize first-strand cDNA using the First Strand cDNA Synthesis Kit (Thermo Fisher Scientific), and real-time RT-PCR was performed with three technical replicates using SYBR Premix Ex Taq (Takara Biotech) on an ABI Applied Biosystems StepOnePlus machine (Life Technologies). Gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method (Kenneth and Schmittgen, 2001). Primers were designed using Primer 5.0, and the primer sequences are listed in **Table 2**.

Data Analysis

In this experiment, the genome of *Vigna angularis* (URL: https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/001/190/045/) was used as a reference genome (Kang et al., 2015). A total of 121.09 G clean data were obtained by sequence similarity analysis by searching against the reference transcriptome database using our 18 samples. The effective data of each sample were 6.04–7.2 G. The distribution of Q30 bases reached 93.7–94.64%, and the average GC content was 44.10%. Mapping reads to the

TABLE 2 | Primers for qRT-PCR.

Order number	Gene ID	Forward primer sequence (5′ → 3′)	Reverse primer sequence (5′ → 3′)
1	ACTIN	F:CTAAGGCTAATCGTGAGAA	R:CGTAAATAGGAACCGTGT
2	LOC108334289	F:CACAATCACAAAGCAGCATCC	R:TTTAGCGAAGTGGCAAGATGA
3	LOC108345251	F:GCCCCTCCTTTTGTTCTCACT	R:CCTCCCACCTTCCATATTGTT
4	LOC108342301	F:TCCTTGACAAATCCAAAACCC	R:GGAAAAGCGAGGAGAAGAAGA
5	LOC108319999	F:TCAGCAGCATCAACAGGACAA	R:GGTCTCAAAGGGGCTATTCAA
6	LOC108320001	F:ATGTCCCTCCATTAGCCTCAG	R:GAACACCCATAGTTCCCCTCA
7	LOC108328360	F:ATCAAAGATGGAGGAACACCC	R:TCCTGCACAATCTACACTTGC
8	LOC108325858	F:TGTTTCAGTTGTTCCTCCGTG	R:GCGTTGCTTGGCTTTTCTAAT
9	LOC108331524	F:AATGGAAAACGGAACTGAGGA	R:CAGAAGGGTGGAAGTGATGCT
10	LOC108344448	F:TCAAAGATAAGACGCAGCACC	R:TGATTGTTAGGGAGGGAGACG
11	LOC108331766	F:AAAAGGGAGGACCAAGAGCAC	R:TGAGTGGCACAACACCTGAAT
12	LOC108328079	F:AAGAAACGCAGAACTTGACCC	R:GCTTGAATGGCAAAGATGAGG
13	LOC108330684	F:AAGAACCCTGGCAGTGTCAAC	R:ATGCCAACATAGCCAATCTCC
14	LOC108340483	F:AACCAATGCCAACCCTCTTAG	R:CAGTGTCCCATCCATAGTCCC

TABLE 3 | Growth of adzuki bean under short-day photoperiod inducement.

Determination stages	Treatment	Plant height/cm	Stem diameter/cm	Main stem nodes number	Leaf area/cm^2	Shoot dry weight/g
Flowering	CK	27.67 ± 3.93a	0.58 ± 0.06a	14.00 ± 1.00a	37.24 ± 3.90a	8.90 ± 1.23a
	SD-5d	19.60 ± 1.91b	0.47 ± 0.08b	13.67 ± 2.08a	29.68 ± 6.47a	6.95 ± 3.04ab
	SD-10d	19.13 ± 2.58b	0.45 ± 0.11b	12.67 ± 0.58a	26.13 ± 1.21ab	4.30 ± 1.00b
	SD-15d	18.93 ± 2.89b	0.41 ± 0.03b	12.00 ± 1.73a	22.66 ± 4.80b	4.59 ± 2.18b
Podding	CK	33.80 ± 4.10a	0.68 ± 0.12a	18.00 ± 1.00a	51.40 ± 5.22a	18.46 ± 2.37a
	SD-5d	32.10 ± 5.52a	0.64 ± 0.13a	17.33 ± 2.08ab	48.51 ± 1.75a	17.00 ± 3.53a
	SD-10d	30.10 ± 3.50a	0.60 ± 0.07a	16.00 ± 1.73b	45.71 ± 6.89ab	16.23 ± 2.90a
	SD-15d	28.53 ± 2.84a	0.49 ± 0.03b	15.33 ± 0.58b	42.40 ± 2.03b	9.87 ± 1.11a
Seed-fling	CK	41.20 ± 5.90a	0.74 ± 0.11a	20.33 ± 4.04a	55.69 ± 7.70a	52.24 ± 4.70a
	SD-5d	40.37 ± 2.71a	0.73 ± 0.05a	19.00 ± 1.73a	52.92 ± 3.78a	40.87 ± 3.88ab
	SD-10d	37.37 ± 6.06a	0.68 ± 0.04ab	17.00 ± 1.00ab	51.41 ± 8.22a	35.20 ± 4.86ab
	SD-15d	29.97 ± 2.98b	0.62 ± 0.03b	15.33 ± 1.15b	47.59 ± 5.33a	23.05 ± 3.13b

Values are means ± S.E. The different small letters in the same column indicate statistical significance at 0.05 level by DMRT.

reference genome to obtain genome alignment of each sample with alignment rate were 97.09–98.22%. The resulting raw data were quality-filtered using NGS QC Toolkit software to obtain high-quality clean reads. Then, the clean reads were aligned using HISAT2 to the reference genome of species by the genome alignment rate. Gene FPKM expression values were quantified using Cufflinks software. Duncan ANOVA and multiple comparisons were carried out at 0.05 level using SPSS 17.0 software, and images were visualized by BS500/BS500-TR biomicroscopy and PhotoshopCS6 software treatment.

RESULTS

Plant Growth

The plant height, stem diameter, internode number, leaf area, and aboveground dry weight of adzuki bean were all inhibited under short-day inducement conditions (**Table 3**). The longer the short-day induction time, the more the decrease in growth index and the earlier the maturity stage. At the flowering stage, compared with CK, the leaf area decreased significantly under all short-day treatments and plant height decreased significantly under SD-15d. At the podding and seed-filling stages, similar results were observed. All the growth indexes had the largest decrease under SD-15d compared with other treatments.

Flower Bud Differentiation Process and Flowering Characteristics

As seen in **Figure 1**, short-day photoperiod inducement significantly affected the process of floral bud differentiation and the flowering characteristics of adzuki bean. After short-day photoperiod inducement, the differentiation of apical flower buds under different treatments was observed. The findings showed that short-day photoperiod inducement significantly promoted the process of flower bud differentiation in adzuki bean. When the control was at the inflorescence primordium differentiation phase, the apical flower buds of adzuki bean plants under short-day light for 5 days, 10 days, and 15 days had

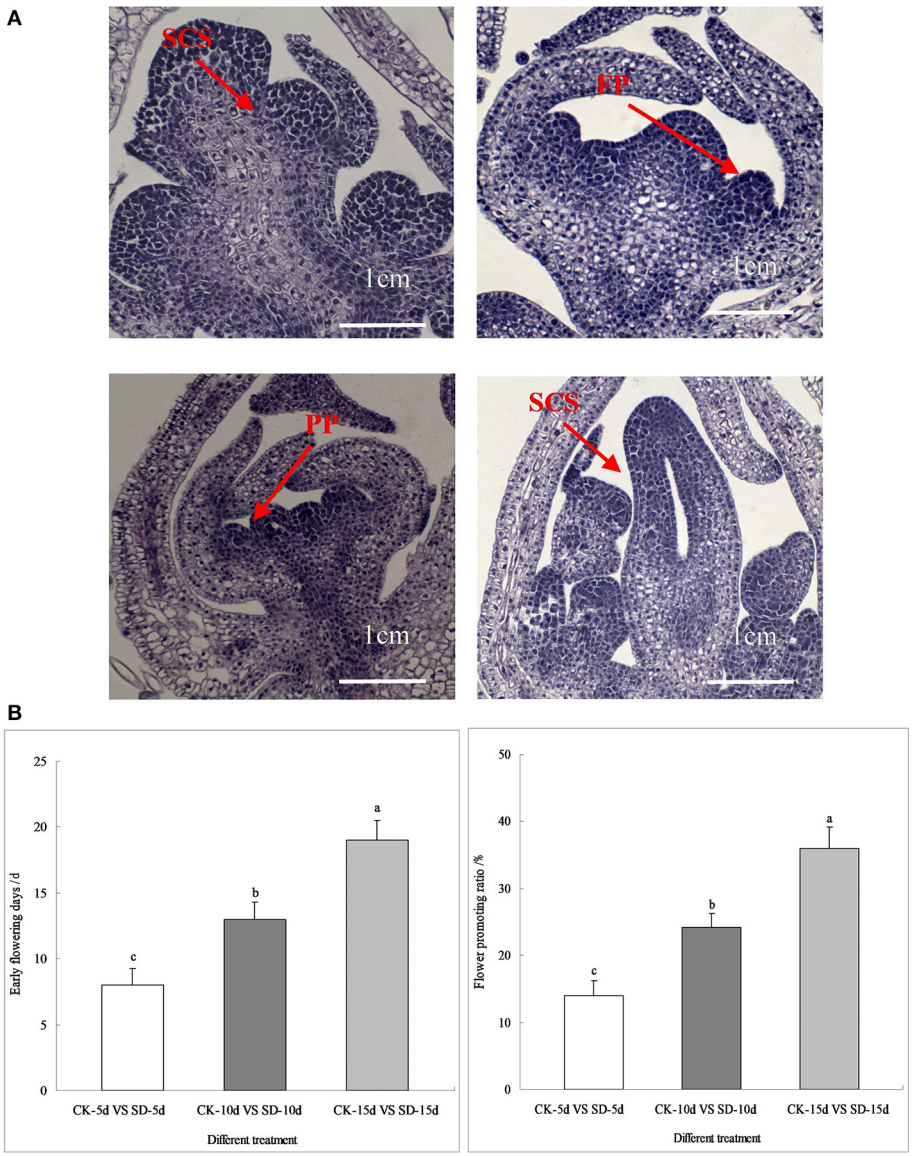


FIGURE 1 | Flower bud differentiation, early flowering days and flower promoting ratio of adzuki bean. **(A)** Pictures of flower bud differentiation processes under different short-day photoperiod treatments. There were three treatments (CK-natural lighting, SD-5d, SD-10d, SD-15d). **(B)** Statistics of early flowering days and flower promoting ratio under different short-day photoperiod treatments (CK-5d vs. SD-5d, CK-10d vs. SD-10d, CK-10d vs. SD-10d). Different letters were represented as Values are means ± S.E. The different lowercase letters above each bar chart indicate statistical significance at 0.05 level by DMRT.

already entered the floral primordium differentiation stage, the petal primordium differentiation stage, and the pistil and stamen differentiation stage (**Figure 1A**).

Further statistics showed that, compared with the controls, adzuki bean with short-day photoperiod inducement of 5, 10, and 15 days had early flowering induced at 7, 13, and 20 days and a flowering promotion rate of 13.99, 24.19, and 36.02%, respectively. The number of days of induced early flowering and the flowering promotion rate under different short-day photoperiod inducement periods reached significant levels. Comparing CK-15d vs. SD-15d to CK-5d vs. SD-5d and CK-10d vs. SD-10d, the number of days of early flowering was 65.00% and 35.00% earlier, respectively, while the flowering promotion rate was 61.16% and 32.84% earlier, respectively (**Figure 1B**). It was seen that the late-maturing "Tang shan hong xiao dou" variety was more sensitive to short-day photoperiod conditions, displaying increased reproductive growth by regulating its own vegetative growth that led to earlier flowering by more number of days with a higher flowering promotion rate. This shows that short-day photoperiod inducement promoted flower bud differentiation and the flowering promotion rate of adzuki bean, possibly due to its role in regulating gene expression in flowering metabolic pathways.

TABLE 4 | Summary of RNA-seq results.

Sample	Raw reads	Clean reads	Q30 (%)	GC content (%)	Total mapped	Multiple mapped	Uniquely mapped
CK-5d-1	46.03M	45.11M	93.70	43.39	44,104,797 (97.77%)	1,944,288 (4.31%)	42,160,509 (93.46%)
CK-5d-2	44.66M	43.84M	94.53	42.54	42,569,240 (97.09%)	1,979,914 (4.52%)	40,589,326 (92.58%)
CK-5d-3	47.70M	46.82M	94.64	42.88	45,654,339 (97.50%)	2,196,016 (4.69%)	43,458,323 (92.81%)
SD-5d-1	45.07M	44.28M	94.57	43.03	43,067,558 (97.27%)	1,975,957 (4.46%)	41,091,601 (92.81%)
SD-5d-2	49.65M	48.65M	93.86	43.98	47,584,894 (97.82%)	2,227,964 (4.58%)	45,356,930 (93.24%)
SD-5d-3	42.68M	41.93M	94.39	42.72	40,719,836 (97.11%)	1,845,739 (4.40%)	38,874,097 (92.71%)
CK-10d-1	50.31M	49.31M	94.03	44.33	48,196,864 (97.75%)	2,325,921 (4.72%)	45,870,943 (93.03%)
CK-10d-2	47.65M	46.71M	94.10	44.75	45,724,496 (97.89%)	2,676,489 (5.73%)	43,048,007 (92.16%)
CK-10d-3	47.91M	47.04M	94.16	44.53	46,153,328 (98.11%)	2,590,091 (5.51%)	43,563,237 (92.61%)
SD-10d-1	43.72M	42.89M	94.04	44.79	42,115,641 (98.19%)	2,619,838 (6.11%)	39,495,803 (92.08%)
SD-10d-2	43.37M	42.54M	93.99	44.43	41,766,075 (98.17%)	2,390,006 (5.62%)	39,376,069 (92.56%)
SD-10d-3	50.18M	49.28M	94.24	44.79	48,403,476 (98.22%)	2,844,589 (5.77%)	45,558,887 (92.45%)
CK-15d-1	50.07M	49.09M	94.00	44.58	48,197,147 (98.18%)	2,475,315 (5.04%)	45,721,832 (93.14%)
CK-15d-2	46.03M	45.21M	94.17	44.29	44,392,867 (98.18%)	2,242,198 (4.96%)	42,150,669 (93.22%)
CK-15d-3	47.81M	46.85M	93.97	44.37	45,866,501 (97.90%)	2,196,580 (4.69%)	43,669,921 (93.21%)
SD-15d-1	49.13M	48.36M	94.42	44.99	47,471,908 (98.16%)	3,039,632 (6.29%)	44,432,276 (91.88%)
SD-15d-2	47.47M	46.69M	94.45	44.62	45,812,022 (98.13%)	2,686,977 (5.76%)	43,125,045 (92.37%)
SD-15d-3	48.25M	47.36M	94.13	44.76	46,500,053 (98.19%)	2,949,831 (6.23%)	43,550,222 (91.96%)

Statistical Analysis and Quality Evaluation of RNA-Seq

The "Tang shan hong xiao dou" variety sampled at 5 days, 10 days, and 15 days of short-day photoperiod treatment was subjected to transcriptome sequencing, and 121.09 G clean data were obtained. The distribution of the proportions of the various reads from the library is given in **Table 4**. The effective data volume of each sample were distributed in the range of 6.04–7.2G, and the distribution of Q30 bases was 93.7–94.64%, with an average GC content of 44.10%. The data obtained from filtering clean reads were compared with those of the reference genome using HISAT2. The degree of matching between the data of each sample (clean reads) and the reference genome sequence was above 90%, with the highest match at 93.46% and the lowest at 91.88%, indicating that the results obtained by sequencing were accurate and reliable.

Numerical Analysis of Differentially Expressed Genes (DEGs) in Response to Short-Day Conditions

Principal component analysis (PCA) results showed that the distance among different samples was long; only CK-5d and SD-5d groups were overlapped. For SD-5d, between-group differences in gene expression were not significant, but others showed larger differences. Although the distance was short within different samples, the gene expression similarity within samples was high, and the gene expression of the samples within the group was uniform and consistent (**Figure 2A**).

The data from the three groups of short-day photoperiod treatments were compared with those from their respective controls, and the results are shown in **Figure 2B**. The results showed that the total number of differential genes was 7,547, of which 4,116 genes were upregulated and 3,431 genes were downregulated in expression. Among them, 88 differential genes were detected between CK-5d vs. SD-5d, including 64 genes that had upregulated expression and 24 genes that had downregulated expression; 2,528 differential genes were detected between CK-10d vs. SD-10d, including 1,222 genes that had upregulated expression and 1,306 genes that had downregulated expression. The highest number of differential genes was detected between CK-15d vs. SD-15d, with a total of 4,931 differential genes, including 2,830 genes that had upregulated expression and 2,101 genes that had downregulated expression. This indicates that the number of differential genes increased with extended durations of short-day photoperiod conditions and the number of flowering response genes increased after short-day photoperiod inducement. As seen in **Figure 2C**, the numbers of identical genes in the three groups were 10, 1,904, and 30, respectively, and there were five differential genes in the three groups, with the largest number of differential genes shared between CK-10d vs. SD-10d and CK-15d vs. SD-15d treatments.

The clustering of differential genes can be seen from the heat map in **Figure 2D**. The gene expression of CK-5d is similar to that of SD-5d, with the lowly expressed genes distributed in the upper and lower parts and the highly expressed genes distributed in the middle. The expression of differential genes changed significantly with prolonged shading time. The highly expressed genes of CK-10d and CK-15d distributed in the lower part and the lowly expressed genes distributed in the upper part. SD-10d and SD-15d treatments showed opposite trends. The aforementioned gene expression results showed that gene expression changed accordingly with prolonged short-day photoperiod treatments, and some genes regulating flowering metabolism such as circadian rhythm, plant hormones, antenna proteins, and photosystem II were upregulated. With

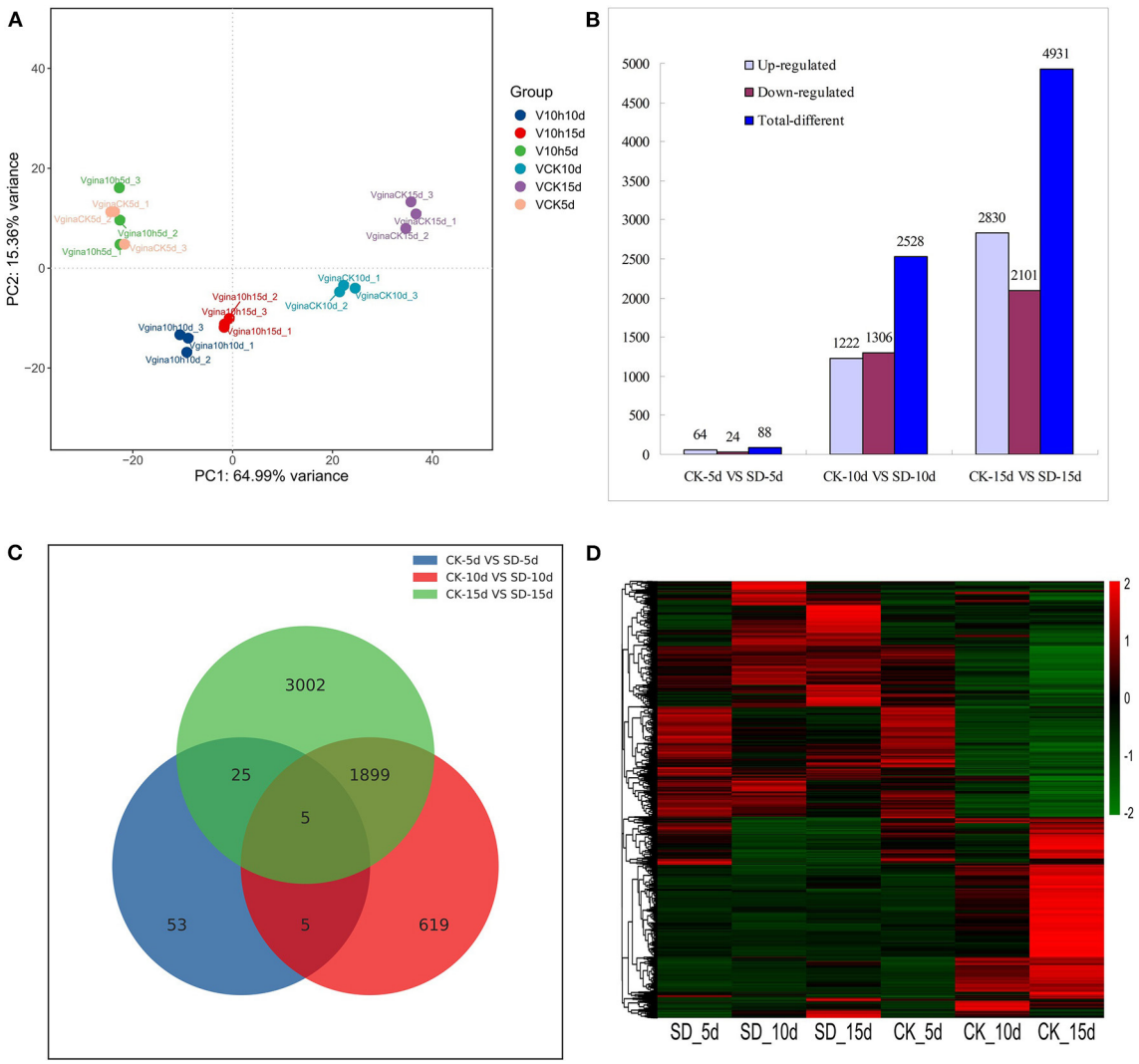


FIGURE 2 | Analysis of differentially expressed genes in response to short-day photoperiod. **(A)** Principal component analysis of the sequenced samples. **(B)** Numbers of the significantly regulated genes under different days short-day photoperiod treatment. The numbers on the horizontal axis reflect different treatment and the vertical axis represents numbers of total, up-and down-regulated genes, respectively. **(C)** Venn diagram analysis of significantly regulated genes under different days short-day photoperiod treatment. **(D)** Heatmap clustering of global pattern of the strongly regulated genes conducted using Hierarchical Clustering (HCL) algorithm under different days short-day photoperiod treatment. The color scale represents the values of lg FPKM (FPKM–Fragments Per Kilobase of transcript per Million fragments mapped.

different response abilities to short-day photoperiod conditions, some genes in other metabolic pathways were downregulated, indicating that the opposite changing trends of gene expression under CK-10d and SD-10d and under CK-15d and SD-15d were related to the earlier flowering of adzuki bean induced by short day.

GO Functional Enrichment Analysis of DEGs in Response to Short-Day Photoperiod Conditions

GO functional enrichment analysis was performed on three groups of differential genes; 10 GO entries corresponding to a number >2 of the differential genes were each screened for biological processes, cellular components, and molecular functions. The GO functional enrichment results (**Figure 3**) showed that each group was significantly different and had different light-related functions according to their corresponding −log10 *p*-value in the descending order, among which CK-5d vs. SD-5d differential genes had a functional group related to flowering, which contained two genes that were both upregulated (**Figure 3A**). CK-10d vs. SD-10d differential genes also had a light-related circadian functional group in the biological processes and contained 42 genes, of which 28 genes were upregulated and 14 genes were downregulated (**Figure 3B**). CK-15d vs. SD-15d differential genes had significant differential

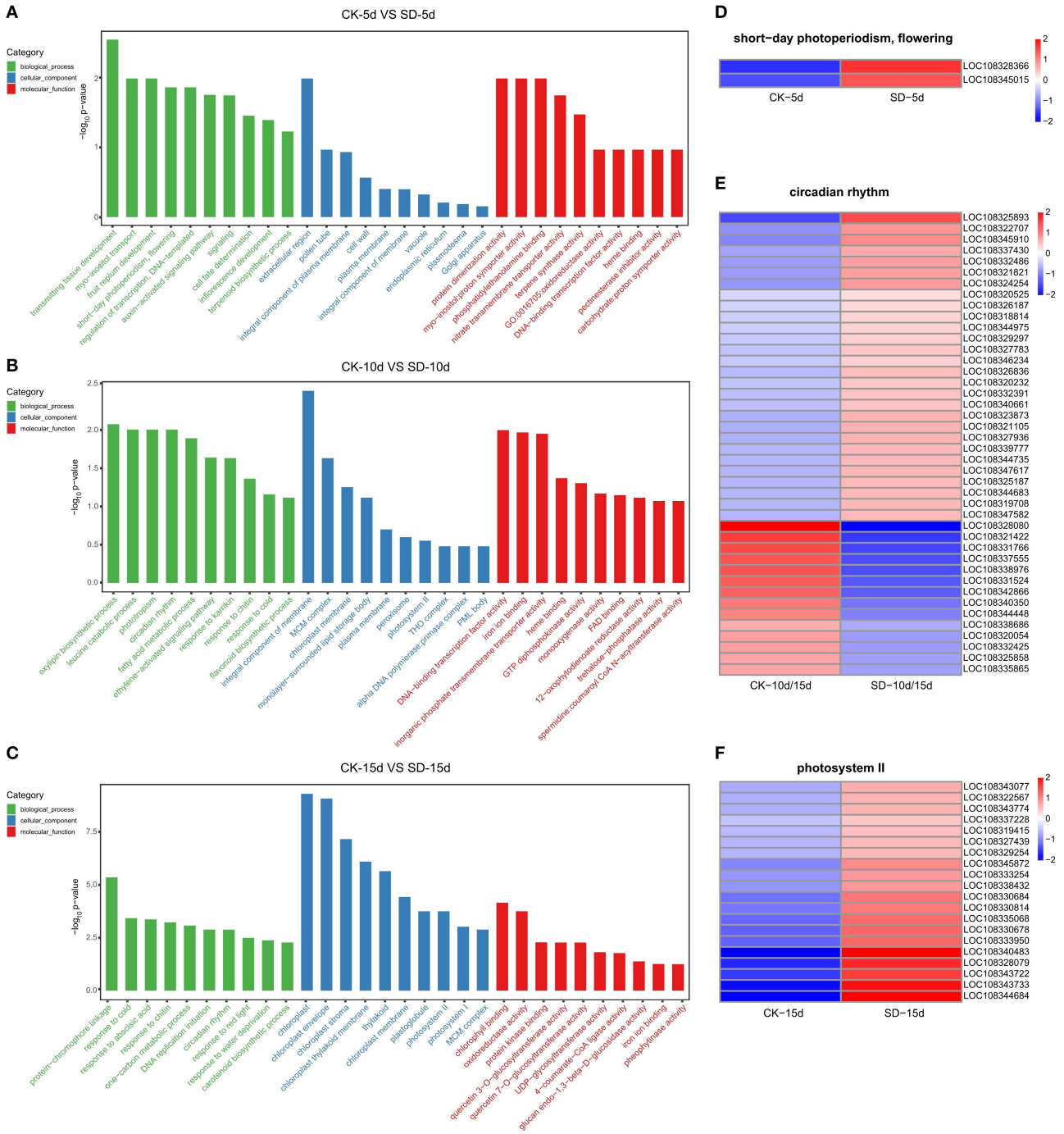


FIGURE 3 | GO enrichment results and heat map clustering analysis of three comparison groups. **(A–C)** GO enrichment between the CK-5d vs. SD-5d, CK-10d vs. SD-10d and CK-15d vs. SD-15d groups, The X coordinates in the diagram are GO entry names, and the Y coordinates are-log10 *p*-values. **(D–F)** Heatmap clustering analysis of short-day photoperiodism, flowering, circadian rhythm and photosystemII functional entries for differential genes. The color scale represents the values of lg FPKM (FPKM–Fragments Per Kilobase of transcript per Million fragments mapped.

expression, coding for the photosystem II functional group in the cell composition related to light in adzuki bean, which contained 20 genes that were all upregulated. Fron the results of the aforementioned analysis, we could speculate that short-day photoperiod inducement can promote flowering by stimulating the upregulation of flowering genes and adjusting the biological clock by upregulating and downregulating genes in the circadian rhythm, thus controlling the flowering time of adzuki bean. In

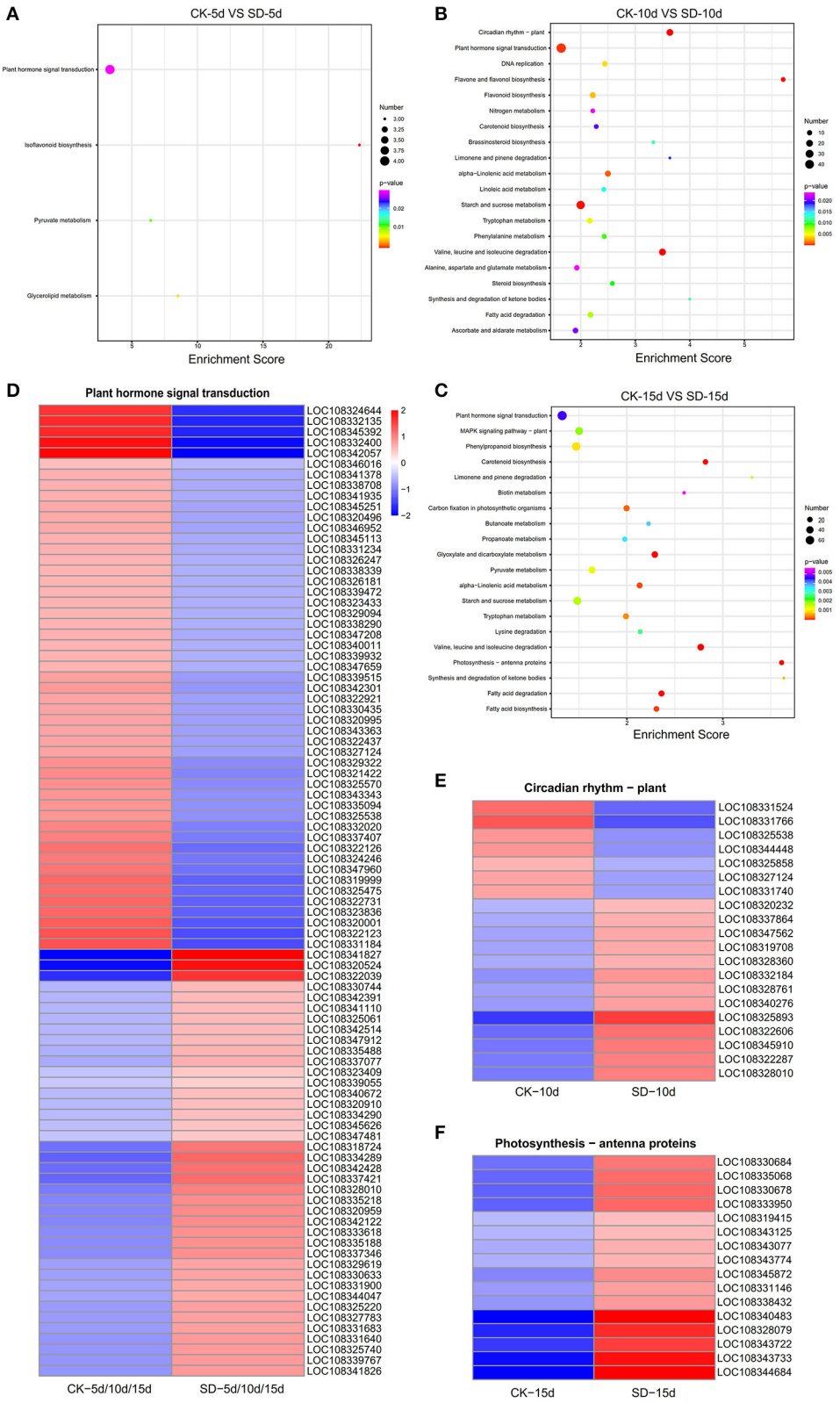


FIGURE 4 | Heat map clustering analysis of KEGG enrichment results of metabolic pathways in the three comparison groups. **(A–C)** KEGG enrichment results of metabolic pathways in CK-5d vs. SD-5d, CK-10d vs. SD-10d, CK-15d vs. SD-15d. Notes: the bubble size represents the number of genes that contain the difference. The number of differentially expressed genes in the pathway with large bubbles was larger. The color of the bubbles represents the degree of enrichment,

(Continued)

FIGURE 4 | i.e., the pvalue was becoming smaller with the significant degree of enrichment becoming greater ranging from purple to red. **(D–F)** Heat map clustering analysis of differential genes in plant hormone signaling pathways, circadian rhythms and antenna protein metabolism pathways. The color scale represents the values of lg FPKM (FPKM–Fragments Per Kilobase of transcript per Million fragments mapped.

addition, short-day photoperiod conditions activate key genes in the photosystem to enhance chlorophyll synthesis and provide nutrients for adzuki bean flowering. Under different short-day photoperiod treatment times, entries were enriched by the GO function, and the expression of genes differed.

Analysis of the KEGG Enrichment Pathway of DEGs in Response to Short-Day Photoperiod Conditions

Figure 4 shows the top 20 pathways with the smallest Q values. The number of the DEGs is indicated by the size of the dots, and the different Q ranges are indicated by the color of the dots. The results of KEGG enrichment analysis for each group of the DEGs showed that the hormone signaling metabolic pathway, the circadian rhythm pathway, and the antenna protein pathway were associated with short-day photoperiod conditions, and more genes were enriched in these pathways. CK-5d vs. SD-5d had more genes enriched in the phytohormone signaling metabolic pathway, of which four were upregulated. CK-10d vs. SD-10d had more genes enriched in the phytohormone signaling metabolic pathway and the circadian pathway; 47 genes were enriched in the phytohormone signaling metabolic pathway, of which 19 genes were upregulated and 28 genes were downregulated, and 20 genes were enriched in the circadian pathway, of which 13 genes were upregulated and 7 genes were downregulated. In the bubble plots of KEGG enrichment for CK-15d vs. SD-15d, the phytohormone signaling metabolic pathway and the antenna pathway had the highest number of enriched genes. Out of the 73 genes enriched in the phytohormone signaling metabolic pathway, 31 genes were upregulated and 42 genes were downregulated, and all 16 genes enriched for the antenna pathway were upregulated. It can also be seen that all three groups of DEGs were enriched in the phytohormone signaling metabolic pathway with the highest number of genes, and after combining the three groups of the DEGs, we obtained 91 genes with 40 genes having upregulated expression and 51 genes having downregulated expression. It was shown that the aforementioned DEGs regulated the endogenous hormone levels of adzuki bean by activating or repressing their own expression and thus directly or indirectly regulating the growth and development of adzuki bean under short-day photoperiod conditions. In addition, the circadian pathway and antenna protein pathway were also enriched to a greater extent, indicating that the expression of relevant differential genes in these two pathways was also active under short-day photoperiod inducement.

GO Function and Expression Analysis of Light-Related Differential Genes in the KEGG Pathway

The temporal changes in the GO function and the number of differential genes associated with light in the KEGG pathway were analyzed. It was found that most genes showed an increase with extended durations of short-day photoperiod inducement, but a few genes were slightly different; two genes were expressed at 5 days among the GO functional entries of short-day photoperiodism and flowering, while one gene was expressed at both 10 days and 15 days. The number of genes in both circadian rhythm and photosystem II functional entries increased with extended durations of short-day photoperiod inducement, and the variations were 0, 24, 38 and 0, 8, 20, respectively (**Figure 5A**). This indicates that circadian rhythm and photosystem II functional entries play a dominant role in response to short-day photoperiod inducement. The number of genes in the light-related pathways of plant hormone signal transduction and photosynthesis-antenna proteins increased with extended durations of short-day photoperiod inducement, with variations of 4, 47, 73, 0, 5, and 16, while the number changes of genes in the circadian rhythm pathway showed an increase and then a slight decrease (**Figure 5B**). Among these three pathways, the largest changes in the number of genes in the phytohormone signal transduction pathway were observed. This indicates that the genes in the metabolic pathway of

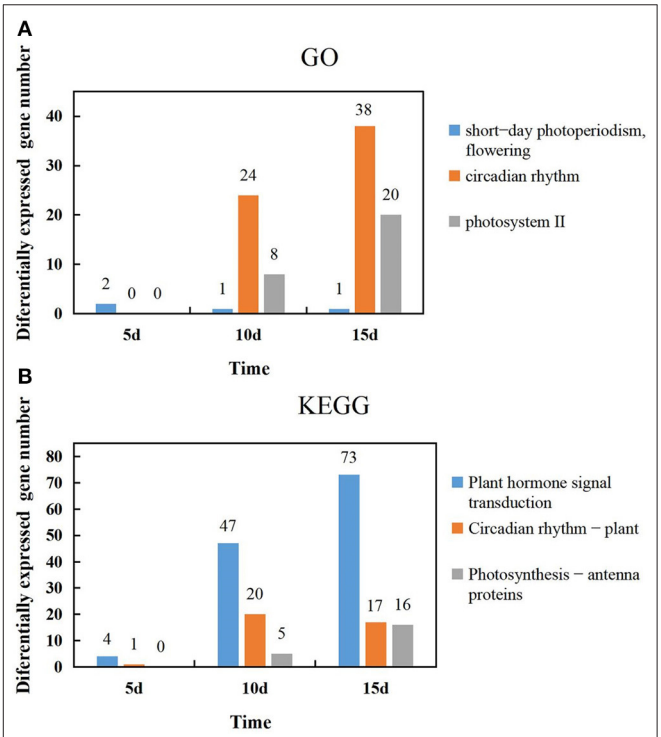


FIGURE 5 | Temporal changes in the number of light-related genes in GO function and the KEGG pathway. **(A)** Changes of the genes number in short-day photoperiodism, flowering, circadian rhythm and photosystem II GO function at 5d, 10d, and 15d. **(B)** Changes of the genes number in Plant hormone signal transduction, Circadian rhythm-plant, Photosynthesis-antenna proteins KEGG pathway at 5d, 10d, and 15d.

phytohormone signaling are coordinated with each other to participate in the control of flowering time in adzuki bean after short-day photoperiod inducement.

In each of the phytohormone signaling metabolic, circadian, and antennal protein signaling pathways, 5, 5, and 3 genes, respectively, that were significantly upregulated or downregulated were selected for comparative expression analysis. As few genes were enriched in KEGG in the CK-5d vs. SD-5d comparison group, genes in each pathway were selected in the CK-10d vs. SD-10d and CK-15d vs. SD-15d comparison groups. The genes selected from both the phytohormone signaling metabolic pathway and the circadian pathway were found to be one gene with upregulated expression and four genes with downregulated expression, while genes in the antenna protein signaling pathway mostly had upregulated expression. A total of 13 differentially expressed genes screened were related to light and flowering according to the homolog function (**Table 5**). It was shown that these genes regulated the growth and development of adzuki bean by activating or repressing their own expression and thus regulating the flowering time of adzuki bean in different metabolic pathways.

qRT-PCR Validation of Differential Gene Expression

To verify the reliability of adzuki bean transcriptome data, 13 genes with significant differences in the plant hormone, circadian rhythm, and antenna protein pathways were verified by qRT-PCR again with new adzuki bean samples using the same treatment. The expression of five genes in the metabolic pathway of phytohormone signaling differed, where

TABLE 5 | List of different genes associated with light exposure in different signaling pathways.

KEGG pathway	Gene ID	gene symbol	Regulation	q-value CK-10dVS SD-10d CK-15dVS SD-15d	Description	Location	Functional annotations of orthologs
Plant hormone signal transduction(KEGG)	LOC108334289	F19P19.31	Up	5.2419E-68 5.1174E-63	Auxin-responsive protein IAA17	Chromosome 5 NC_030641.1	Auxin-responsive protein
	LOC108345251	T20010.80	Down	5.7337E-25 2.9722E-25	Protein transport inhibitor response 1-like	Chromosome 1 NC_030637.1	Leucine-rich repeat, cysteine-containing subtype
	LOC108342301	MJB21.13	Down	1.6981E-17 4.8082E-26	BRI1 kinase inhibitor 1-like	Chromosome 9 NC_030645.1	BRI1 kinase inhibitor 1-like
	LOC108319999	MOE17.8	Down	5.7916E-40 2.5844E-153	Protein ethylene insensitive 3-like	Un NW_016114933.1	Ethylene insensitive 3
	LOC108320001	MOE17.8	Down	1.5701E-35 3.6499E-144	Protein ethylene insensitive 3-like	Un NW_016114933.1	Ethylene insensitive 3
Circadian rhythm-plant (KEGG)	LOC108328360	MKP6.27	Up	1.3414E-18 1.7789E-41	Transcription factor HY5-like	Chromosome 3 NC_030639.1	Basic-leucine zipper domain
	LOC108325858	F19P19.14	Down	2.9319E-16 1.8408E-106	Cryptochrome-1-like	Chromosome 2 NC_030638.1	Cryptochrome/DNA photolyase, FAD-binding domain
	LOC108331524	T3H13.14	Down	1.9432E-53 9.3532E-124	Cryptochrome-1	Chromosome 4 NC_030640.1	Cryptochrome/DNA photolyase, FAD-binding domain
	LOC108344448	F16B22.17	Down	1.4873E-20 1.3516E-97	Putative casein kinase II subunit beta-4	Chromosome 1 NC_030637.1	Casein kinase II subunit beta
	LOC108331766	F17H15.25	Down	1.9738E-62 2.0339E-153	Protein early flowering 3-like	Chromosome 4 NC_030640.1	Protein early flowering 3-like
Photosynthesis-antenna proteins (KEGG)	LOC108328079	F27I1.2	Up	2.3086E-22 1.2264E-122	Chlorophyll a-b binding protein CP29.3, chloroplastic	Chromosome 3 NC_030639.1	Chlorophyll a-b binding protein
	LOC108330684	lhcA-P4	Down/Up	6.1631E-16 1.42167E-33	Chlorophyll a-b binding protein P4, chloroplastic	Chromosome 4 NC_030640.1	Chlorophyll a-b binding protein
	LOC108340483	LHBC1	Up	3.3249E-11 1.8627E-57	Chlorophyll a-b binding protein 13, chloroplastic	Chromosome 8 NC_030644.1	Chlorophyll a-b binding protein

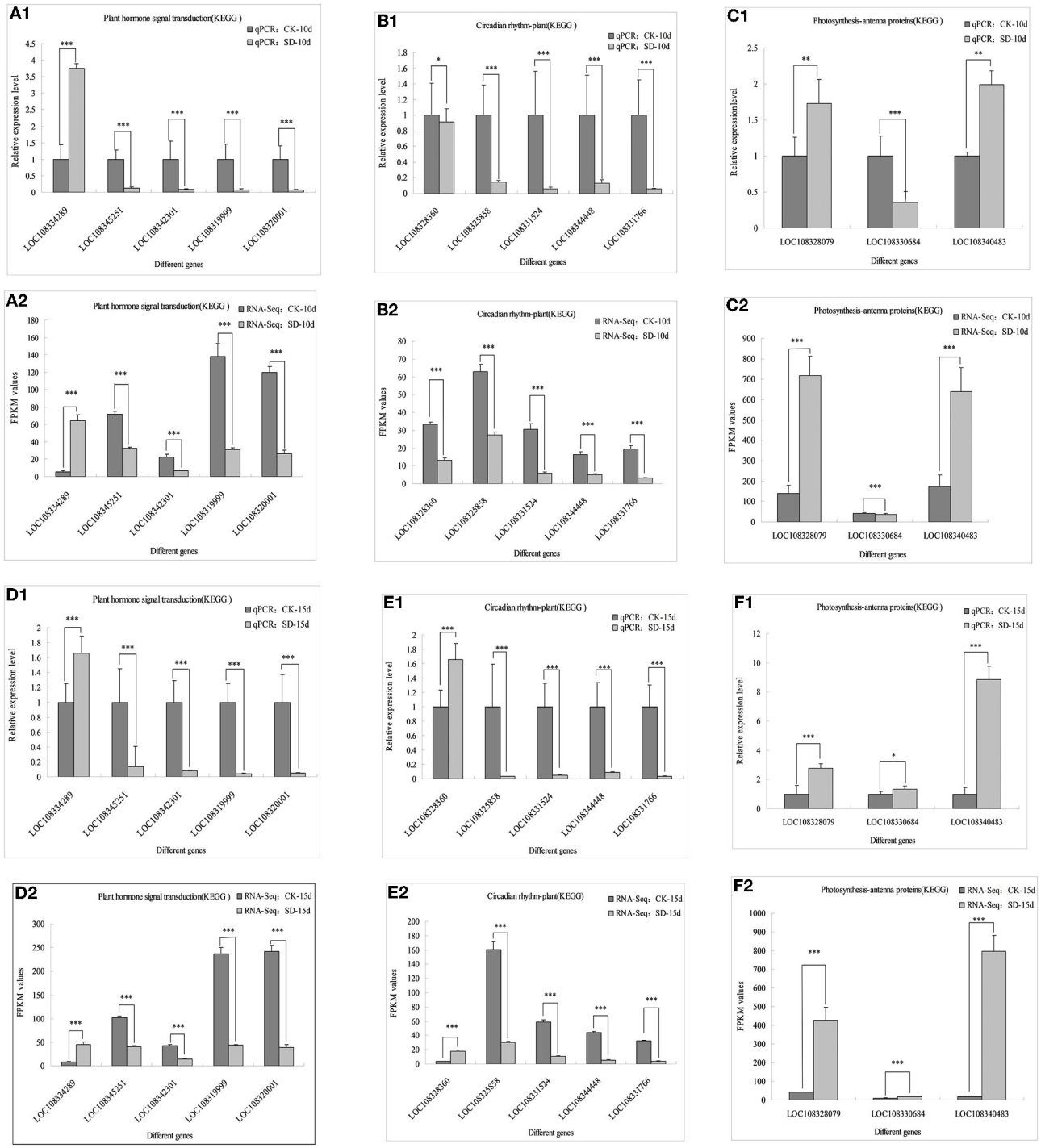


FIGURE 6 | Comparation between qRT-PCR validation and RNA-seq of 13 differentially related photoperiod expressed genes. *Indicates significant difference at 0.05 level, ** and *** indicate extremely significant difference at 0.01 and 0.001 level. **(A1–C1)** The qPCR validation of 13 genes in three metabolic pathways of plant hormone signal transduction, circadian rhythm and photosynthesis-antenna protein were induced for 10 days under 10 h short-day photoperiod inducement; **(A2–C2)** RNA-Seq values of 13 genes in three metabolic pathways of plant hormone signal transduction, circadian rhythm and photosynthesis-antenna protein were induced for 10 days under 10 h short-day photoperiod inducement; **(D1–F1)** The qPCR validation of 13 genes in three metabolic pathways of plant hormone signal transduction, circadian rhythm and photosynthesis-antenna protein were induced for 10 days under 15 h short-day photoperiod inducement; **(D2–F2)** RNA-Seq values of 13 genes in three metabolic pathways of plant hormone signal transduction, circadian rhythm and photosynthesis-antenna protein were induced for 10 days under 15 h short-day photoperiod inducement.

expression of *LOC108334289* significantly increased under SD-10d and SD-15d treatments compared with that in the control, whereas *LOC108345251*, *LOC108342301*, *LOC108319999*, and *LOC108320001* had significantly lower expression than those in the control (**Figures 6A1,D1**). In the circadian rhythm pathway, *LOC108328360* genes had significantly lower expression under SD-10d treatment, but significantly increased expression under SD-15d treatment. Moreover, *LOC108325858*, *LOC108331524*, *LOC108344448*, and *LOC108331766* had significantly lower expression under both SD-10d and SD-15d treatments than in the control (**Figures 6B1,E1**). In the antenna signaling pathway, *LOC108330684* had significantly reduced expression compared with that in the control under SD-10d treatment but significantly increased expression under SD-15d treatment, with *LOC108328079* and *LOC108340483* having significantly elevated expression compared with the control under SD-10d and SD-15d treatments (**Figures 6C1,F1**). In addition, qRT-PCR analysis on detecting expression patterns of 13 genes indicated that they were completely consistent with the transcriptional sequencing results RNA-Seq values (**Figures 6A1A2–F1F2**), and the differences of 13 genes in different signaling pathways reached a significant level. Only *LOC108328360*, *LOC108328079*, and *LOC108340483* genes had different differences under 10-day treatment, while LOC108330684 genes had different differences under 15-day treatment. These results indicated that these transcriptome sequencing data were accurate and reliable.

DISCUSSION

Adzuki bean (*Vigna angularis*) originated from and produced in China, with cropland area and total production as one of the largest in the world. Adzuki bean can be adapted to a wide range of uses and is a grain legume of economic importance. As a model crop with a short growth cycle and small genome, adzuki bean is often used to study model plants through transcriptomics to reveal the biological mechanisms of different genes and has important applications in the genetic improvement of legume crops (Yamada et al., 2001). Currently, transcriptome sequencing technology is widely used in the discovery of new genes and comparative genomics studies (Wang et al., 2014; Prince et al., 2015; Liu et al., 2017; Zhang et al., 2017). In this study, we used 'Tang shan hong xiao dou', which is sensitive to short-day photoperiod conditions, as the material to construct a cDNA library of bean under full-day and short-day photoperiod treatment conditions, and we conducted transcriptome sequencing using Illumina sequencing technology to screen DEGs. DEGs increased with extended durations of short-day induction. Short-day photoperiod inducement of 15 days had the highest number of 4,931 DEGs, of which 2,830 were upregulated and 2,101 were downregulated. In terms of GO functional enrichment of DEGs, in CK-5d vs. SD-5d comparison, the differential genes associated with short-day photoperiod conditions were mostly focused on the biological process of flowering promotion, but there were only two genes for this function and related to flowering time; both are described as protein HEADING DATE 3B-like (Hd3) and protein FLOWERING LOCUS T-like (FT), respectively. Related findings in other species showed that Hd3 is a homolog of FT that can exert its floral function as an anthocyanin gene, with the protein expressed in leaf vascular bundles and transferred to the apical meristem through the vascular bundle to promote plant flowering (Li et al., 2014). It was speculated that adzuki bean was blossomed earlier after 5 days of short-day photoperiod inducement, being related to accumulation and transportation of florigen. Through the biological processes, the comparison of CK-10d vs. SD-10d was dominated by 42 differentially expressed circadian rhythm-related genes, of which 28 genes were upregulated and 14 genes were downregulated. Previous studies on other crops showed that the GI-CO-FT cascade in circadian pathways acts as a regulator in determining the flowering behavior (Nakamichi et al., 2020; Liu et al., 2021; Wang et al., 2021). It was speculated that adzuki bean may also participate in this pathway to promote flowering under short-day photoperiod inducement. The CK-15d vs. SD-15d comparison showed that there were 20 DEGs of PS II in the cellular group, and all these genes reached significant levels and were upregulated. It is indicated that these genes may be related to the antenna protein genes. Peripheral light-capturing antenna of photoperiod system II can capture and transmit light energy, improve the utilization rate of light energy, and promote early flowering of adzuki bean under short-day photoperiod inducement conditions. By comparing GO temporal changes in the number of genes in different functions, the number of genes directly promoting flowering in adzuki bean through short-day photoperiod inducement did not increase with the duration of short-day photoperiod condition treatments but rather decreased. However, PS II and circadian rhythm function-related genes increased with the extension of the short-day photoperiod inducement duration, indicating that the gene related to the function of photoperiod-induced plant flowering did not play a dominant role, while PS II and circadian rhythm function might be the key genes regulating the growth and development of adzuki bean under short-day photoperiod inducement.

The results of the KEGG metabolic pathway analysis revealed that the metabolic pathway of phytohormone signaling associated with short-day photoperiod inducement was enriched with the most DEGs, followed by the circadian pathway, and the antenna protein pathway. This suggests that these three pathways may be critical for short-day photoperiod inducement to interfere with the growth and development of adzuki bean and that these genes may be key factors or candidate genes affecting the regulation of short-day photoperiod inducement in adzuki bean seedlings. This is similar to the findings obtained in previous studies using transcriptome sequencing in other crops under short-day photoperiod inducement (He et al., 2017; Xiao et al., 2017; Gao et al., 2021). In addition, we selected 5, 5, and 3 candidate genes with significant differences in regulating flowering through the phytohormone, circadian, and antenna protein metabolism pathways, respectively, because adzuki bean is most closely associated with the phytohormone signaling pathways under short-day photoperiod inducement, and the phytohormone signaling pathways play an important role in the induction of flower formation in plants. By transcriptome

analysis, 5 candidate genes in the phytohormone metabolic pathway was described as auxin-responsive protein IAA17, transport inhibitor response 1-like protein, BRI1 kinase inhibitor 1-like protein, and ethylene insensitive 3-like protein. Related findings in other species showed that auxin (IAA) promotes female flower development by inducing ethylene synthesis (Song S. S. et al., 2013); transcription of TIR1 is induced by IAA and ABA; auxin can enhance the interaction between TIR1 and (Aux/IAA) (Yun et al., 2011); the BRI1 kinase repressor can inhibit BRI1 by acting in conjunction with other receptor proteins, eliciting BRI1 in an activated state and prompting brassinolactone gene expression to regulate the flowering time in plants (Jia et al., 2020); and the nuclear protein EIN3 binds to the ethylene reaction gene *ERF1* to initiate a series of transcriptional cascades regulating the expression of ethylene target genes (Kosugi and Ohashi, 2000). In this study, short-day photoperiod conditions not only promoted floral bud differentiation in adzuki bean but also advanced flowering time and flowering promotion rate. The analysis of the transcriptome data revealed that the largest numbers of DEGs associated with phytohormone signaling were mainly growth hormone, gibberellin, and ethylene (ETH), suggesting that short-day photoperiod inducement of flowering in adzuki bean was due to the synergistic effect of multiple phytohormones, among which IAA may play a dominant role. This has some similarities with previous studies (Yang et al., 2014; Zhang et al., 2015). These results suggest that the flowering time of adzuki bean may be related to the change in endogenous hormone levels caused by the upregulation/downregulation of DEGs.

Transcriptome analysis revealed that 5 candidate genes involving the circadian rhythm were described as protein early flowering 3-like, cryptochrome-1-like, cryptochrome-1, putative casein kinase II subunit beta-4, and transcription factor HY5-like, respectively. In other species, we found that HY5 indirectly regulates the synthesis and accumulation of anthocyanins in plants by activating the transcription of itself and downstream target genes after receiving signals transmitted by upstream photoreceptors (Zhang et al., 2010). CRY can act with PHY and PHOT proteins to regulate COP1, a downstream factor in the photoreceptor pathway, to modulate the flowering time in plants (Li et al., 2018). Casein kinase 2 (CK2) is involved in many important physiological processes, such as physiological clock, photoperiod, and flower development of plants (Jayant and Enamul, 2012). EFL3 plays an important role in regulating plant early flowering through the circadian clock pathway (Hicks et al., 2001). It has been suggested that the circadian clock system can coordinate external light temperature signals and the endogenous metabolic and developmental state cues, outputting circadian diurnal signals to regulate plant growth and development after short-day inducement; 3 antenna protein candidate genes were descripted as chlorophyll a-b-binding proteins CP29.3, P4, and 13, chloroplastic, respectively. All these genes in *Barley* and *Arabidopsis* have main functions in the capturing and transmitting of light energy, balancing the excitation energy of PS I and PS II by capturing and transmitting light energy in the photosystem and especially in maintaining the structure of the vesicle-like membrane and responding to external environmental changes and photoprotection (Xia Y. S. et al., 2012; Xu et al., 2012). In this study, 16 genes were enriched by KEGG to the antenna protein, including five LHCII-encoded major photopigment protein complexes, all of which were upregulated in expression and associated with photoperiodic responses, suggesting a stress response of adzuki bean after a short-day photoperiod inducement to improve its own light energy utilization to accumulate more compounds to maintain the cellular osmotic potential. These findings were verified in rice, *Arabidopsis* (Pavan, 2010), *Ginkgo* (Wang et al., 2015), and *Mongolian wheatgrass* (Zhao et al., 2017). By analyzing the temporal changes in the number of genes of different metabolic pathways in KEGG, we found that the number of genes in circadian rhythm genes showed a trend of increase and then decrease in the three groups of CK-5d vs. SD-5d, CK-10d vs. SD-10d, and CK-15d vs. SD-15d, while the number of genes in the phytohormone signal transduction metabolic pathway increased with the extension of short daylight, with the largest number of genes in the phytohormone signal transduction metabolic pathway. Thus, the hypothesis is that, after short-day photoperiod inducement, the genes of the phytohormone signal transduction metabolic pathway play a dominant role in growth and development, and flowering is regulated through the synergistic effect of multiple phytohormones. The results of these studies elucidate the molecular regulatory basis of photoperiod-induced flowering in adzuki bean and provide a theoretical basis for photoperiod-adapted breeding of adzuki bean germplasm, as well as lay the foundation for functional validation of DEGs for future research.

CONCLUSION

Comparing the transcriptomes of CK-5d vs. SD-5d, CK-10d vs. SD-10d, and CK-15d vs. SD-15d in the late maturing variety "Tang shan hong xiao dou" indicated that there were 88, 2,528, and 4,931 DEGs, respectively, indicating that the DEGs increased with the extension in the duration of short-day photoperiod inducement. The 15d short-day photoperiod condition was observed to have the greatest effect on flower bud differentiation and the flowering time of adzuki bean. To reveal the expression of genes related to short-day photoperiod inducement and flowering, the analysis of GO and KEGG revealed that the DEGs were mainly associated with functional entries in flowering, photosystem, and circadian rhythm. Pathway analysis of differential protein-encoding genes using the KEGG database further revealed that most genes were mainly enriched in the hormone signaling metabolism, circadian rhythm, and antenna protein pathways; 13 flowering-associated candidate genes were identified in the three pathways, which may be involved in the regulation of flowering time in adzuki bean. The present study provides insights into the elucidation of the gene transcription and regulatory network of flowering time and an empirical basis for further studies of the development and flowering characteristics in adzuki bean species.

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: NCBI - PRJNA817421.

AUTHOR CONTRIBUTIONS

WD and YZ designed the experiments and revised the manuscript. DL and LZ performed the field experiment and collected samples. LZ and BY collected phenotypic data in Baoding in 2021. WD and DL performed data analysis and wrote the manuscript. All authors read and approved the final manuscript version.

FUNDING

This work was supported by the Hebei Province Natural Science Foundation for Youth (C2021204405), the China Agriculture Research System of MOF and MARA – Food Legumes (CARS-08-G-22), and the National Key Research and Development Program of China (2021YFD1901004-2).

ACKNOWLEDGMENTS

The authors thank the reviewers for their valuable comments on this manuscript and other students for their assistance during the preparation of this manuscript.

REFERENCES

Anders, S., and Huber, W. (2012). Differential expression of RNA-Seq data at the gene level-the DESeq package. *Eur. Mol. Biol. Lab*. 23–25. Available online at: http://bioconductor.riken.jp/packages/3.4/bioc/vignettes/DESeq/inst/doc/DESeq.pdf

Banfield, M. J., Barker, J. J., Perry, A. C., and Brady, R. L. (1998). Function from structure? The crystal structure of human phosphatidylethanolamine-binding protein suggests a role in membrane signal transduction. *Structure* 6, 1245–1254. doi: 10.1016/S0969-2126(98)00125-7

Basu, U., Narnoliya, L., Srivastava, R., Sharma, A., Bajaj, D., Daware, A., et al. (2019). Clavata signaling pathway genes modulating flowering time and flower number in chickpea. *Theor. App. Genet.* 132, 2017–2038. doi: 10.1007/s00122-019-03335-y

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

Cheng, X., Li, G., Tang, Y., and Wen, J. (2018). Dissection of genetic regulation of compound inforescence development in *Medicago truncatula*. *Development* 145, 158766. doi: 10.1242/dev.158766

Cober, E. R., Molnar, S. J., Charette, M., and Voldeng, H. D. (2010). A new locus for early maturity soybean, *Crop Sci.* 50, 524–527. doi: 10.2135/cropsci2009.04.0174

Cober, E. R., and Voldeng, H. D. (2001). A new soybean maturity and photoperiod sensitivity locus linked to and photoperiod-sensitivity locus linked to E1 and T. *Crop Sci.* 41, 698–701. doi: 10.2135/cropsci2001.413698x

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., et al. (2007). FT protein movement contributes to long-distance signaling in foral induction of *Arabidopsis*. *Science* 316, 1030–1033. doi: 10.1126/science.1141752

Dissanayaka, A., Rodriguez, T. O., Di, S., Yan, F., Githiri, S. M., Rodas, F. R., et al. (2016). Quantitative trait locus mapping of soybean maturity gene *E5*. *Breed. Sci*. 66, 407–415. doi: 10.1270/jsbbs.15160

Dong, W. X., Zhang, Y. Y., Zhang, Y. L., Ren, S., Wei, Y., and Zhang, Y. C. (2016). Short-day photoperiod effects on plant growth, flower bud differentiation, and yield formation in adzuki bean (*Vigna angularis* L.). *Int. J. Agric. Biol.* 18, 337–345. doi: 10.17957/IJAB/15.0091

Fang, C., Liu, J., Zhang, T., Su, T., Li, S. C., Cheng, Q., et al. (2021). A recent retrotransposon insertion of J caused E6 locus facilitating soybean adaptation into low latitude. *J. Integr. Plant Biol.* 63, 995–1003. doi: 10.1111/jipb.13034

Gao, Y., Wang, X. L., Tu, L., Liu, P. F., Guo, X. Y., Wang, A. G., et al. (2021). Transcriptional level difference analysis between tropical maize inbred lines under two different lighting conditions. *Mol. Plant Breeding* 19, 6972–6983. doi: 10.13271/j.mpb.019.006972

He, P., Li, L. G., Wang, H. B., Chang, Y. S., and Li, H. F. (2017). Effects of shading fruit with opaque paper bag on transcriptome in peach. *Sci Agric. Sinica* 50, 1088–1097. doi: 10.3864/j.issn.0578-1752.2017.06.010

Hicks, K. A., Albertson, T. M., and Wagner, D. R. (2001). *EARLY FLOWERING3* encodes a novel protein that regulates circadian clock function and flowering in *Arabidopsis*. *Plant Cell* 13, 1281–1292. doi: 10.1105/TPC.010070

Izawa, T. (2007). Daylength measurements by rice plants in photoperiodic short-day flowering. *Int. Rev. Cytol.* 256, 191–222. doi: 10.1016/S0074-7696(07)56006-7

Jaudal, M., Zhang, L. L., Che, C., Hurley, D. G., Thomson, G., Wen, J., et al. (2016). MtVRN2 is a Polycomb VRN2-like gene which represses the transition to flowering in the model legume *Medicago truncatula*. *Plant J.* 86, 145–160. doi: 10.1111/tpj.13156

Jayant, M. J., and Enamul, H. (2012). Does CK2 affect flowering time by modulating the autonomous pathway in *Arabidopsis*? *Plant Signal. Behav.* 7, 292–294. doi: 10.4161/psb.18883

Jeong, S., and Clark, S. E. (2005). Photoperiod regulates flower meristem development in *Arabidopsis thaliana*. *Genetics* 169, 907–915. doi: 10.1534/genetics.104.033357

Jia, D. D., Chen, L. G., Yin, G. M., Yang, X. R., Gao, Z. H., Guo, Y., et al. (2020). *Brassinosteroids* regulate outer ovule integument growth in part via the control of INNER NO OUTER by BRASSINOZOLE-RESISTANT family transcription factors. *J. Integr. Plant Biol.* 62, 1093–1111. doi: 10.1111/jipb.12915

Jin, W. L., and Wang, Q. Y. (1995). Study on differentiation of adzuki bean buds. *J. Nanjing Agric. Univ*. 18, 15–20.

Kang, Y. J., Satyawan, D., Shim, S., Lee, T., Lee, J., Hwang, W. J., et al. (2015). Draft genome sequence of adzuki bean, *Vigna angularis*. *Sci. Rep.* 5, 8069. doi: 10.1038/srep08069

Kenneth, 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.2001.1262

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

Kong, D. Y., Chen, S. J., Zhou, L. G., Gao, H., Luo, L. J., and Liu, Z. C. (2016). Research progress of photoperiod regulated genes on flowering time in rice. *Hereditas* 38, 532–542. doi: 10.16288/j.yczz.15-478

Kong, F. J., Liu, B. H., Xia, Z. J., Sato, S., Bo, M. K., Watanabe, S., et al. (2010). Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* 154, 1220–1231. doi: 10.1104/pp.110.160796

Kosugi, S., and Ohashi, Y. (2000). Cloning and DNA-bing properties of a tobacco ethylene- Insensitive3 (EIN3) homolog. *Nucl. Acids Res.* 28, 960–967. doi: 10.1093/nar/28.4.960

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, G. X., Chen, A. H., Liu, X., Wang, W. Y., Ding, H. F., Li, J., et al. (2014). QTL detection and epistasis analysis for heading date using single segment substitution lines in rice (*Oryza sativa* L.). *J Integr. Agr.* 13, 2311–2321. doi: 10.1016/S2095-3119(13)60615-2

Li, S. M., Zhou, Z., Tu, M., Ye, X. X., Lin, C. T., et al. (2018). Research progress of photo signal pathway of cryptochrome in Arabidopsis thaliana. *Mol. Plant Breed*. 16, 4444–4452. doi: 10.13271/j.mpb.016.004444

Lin, X., Liu, B., Weller James, L., Abe, J., and Kong, F. (2021). Molecular mechanisms for the photoperiodic regulation of flowering in soybean. *J. Integr. Plant Biol.* 63, 981–994. doi: 10.1111/jipb.13021

Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K., and Abe, J. (2008). Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome a gene. *Genetics* 180, 995–1007. doi: 10.1534/genetics.108.092742

Liu, D. Q., Jin, Y. H., Guo, L., Li, Y. G., and Huang, X. Z. (2021). Identification and evolutionary analysis of CONSTANS-like gene family in four cruciferous plants. *Plant Physiol. J.* 57, 1241–1260. doi: 10.13592/j.cnki.ppj.2021.0003

Liu, H. B., Liu, X. L., Su, H. S., Lu, X., Xu, C. H., Mao, J., et al. (2017). Transcriptome difference analysis of saccharum spontaneum roots in response to drought stress. *Sci. Agric. Sinica* 50, 1167–1178.

Liu, M., Jiang, X. M., and Zhang, Y. C. (2015). Effect of adzuki bean immature flowers' differentiating process on the law of flower pod formation and abscission. *J. Henan Agri. Sci.* 44, 39–42. doi: 10.15933/j.cnki.1004-3268.2015.07.010

Liu, W., Jiang, B. J., Ma, L. M., Zhang, S. W., Zhai, H., Xu, X., et al. (2018). Functional diversification of *Flowering Locus T* homologs in soybean: *Gmft1a* and *Gmft2a/5a* have opposite roles in controlling flowering and maturation. *New Phytol.* 217, 1335–1345. doi: 10.1111/nph.14884

Lu, S. J., Zhao, X. H., Hu, Y. L., Liu, S. L., Nan, H. Y., Li, X. M., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49, 773–779. doi: 10.1038/ng.3819

Nakamichi, N., Tudo, T., Makita, N., Kiba, T., Kinoshita, T., and Sakakibara, H. (2020). Flowering time control in rice by introducing *Arabidopsis* clock-associated PSEUDO-RESPONSE REGULATOR 5. *Biosci. Biotech. Bioch.* 84, 970–979. doi: 10.1080/09168451.2020.1719822

Pajoro, A., Biewers, S., Dougali, E., Valentim, F. L., Mendes, M. A., Porri, A., et al. (2014). The evolution of gene regulatory networks controlling Arabidopsis plant reproduction: a two-decade history. *J. Exp. Bot.* 65, 4731–4745. doi: 10.1093/jxb/eru233

Pavan, U. (2010). Genome-wide analysis of the family of light-harvesting chlorophyll a/b-binding proteins in *Arabidopsis* and rice. *Plant Signal. Behav.* 5, 1537–1542. doi: 10.4161/psb.5.12.13410

Prince, S. J., Joshi, T., Mutava, R. N., Syed, N., Vitor, M. D. S. J., Patil, G., et al. (2015). Comparative analysis of the drought-responsive transcriptome in soybean lines contrasting for canopy wilting. *Plant Sci.* 240, 65–78. doi: 10.1016/j.plantsci.2015.08.017

Roberts, A., and Pachter, L. (2013). Streaming fragment assignment for real-time analysis of sequencing experiments. *Nat. Methods* 10, 71–73. doi: 10.1038/nmeth.2251

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J., and Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.* 12, R22. doi: 10.1186/gb-2011-12-3-r22

Samanfar, B., Molnar, S. J., Charette, M., Schoenrock, A., Dehne, F., and Golshani, A. (2017). Mapping and identification of a potential candidate gene for a novel maturity locus, E10, in soybean. *Genetics* 130, 377–390. doi: 10.1007/s00122-016-2819-7

Shrestha, R., Gómez-Ariza, J., Brambilla, V., and Fornara, F. (2014). Molecular control of seasonal flowering in rice. *Arabidopsis* and temperate cereals. *Ann. Bot.* 114, 1445–1458. doi: 10.1093/aob/mcu032

Song, S. S., Qi, T. C., Huang, H., and Xie, D. X. (2013). Regulation of stamen development by coordinated actions of jasmonate, auxin, and gibberellin in *Arabidopsis*. *Mol. Plant* 6, 1065–1073. doi: 10.1093/mp/sst054

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, H. L., Choi, C. W., Park, K. M., Jung, W. H., and Min, C. K. (2021). Diversification in functions and expressions of soybean flowering locus t genes fine-tunes seasonal flowering. *Front Plant Sci.* 12, 613675. doi: 10.3389/fpls.2021.613675

Taoka, K., Ohki, I., Tsuji, H., Furuita, K., Hayashi, K., and Yanase, T. (2011). 14–3-3 proteins act as intracellular receptors for rice Hd3a forigen. *Nature* 476, 332–335. doi: 10.1038/nature10272

Tribhuvan, K. U., Das, A., Srivastava, H., Kumar, K., Durgesh, K., Gaikwad, K., et al. (2020). Identifcation and characterization of PEBP family genes reveal CcFT8 a probable candidate for photoperiod insensitivity in *C. Cajan*. *3 Biotech.* 10, 194. doi: 10.1007/s13205-020-02180-x

Wang, F., Zang, X. S., Kabir, M. R., Liu, K. L., Liu, Z. S., Ni, Z. F., et al. (2014). A wheat lipid transfer protein 3 could enhance the basal thermotolerance and oxidative stress resistance of *Arabidopsis*. *Gene* 550, 18–26. doi: 10.1016/j.gene.2014.08.007

Wang, H. L., Liu, X. L., Yu, W. W., Li, G. P., and Cao, F. L. (2015). Cloning of light-harvesting cholorophyll a/b-binding protein coding gene (*GbLhcb4*) and promoter sequence from Ginkgo biloba. *J. Cent. South Univ. Forestry Technol.* 35, 114–121. doi: 10.14067/j.cnki.1673-923x.2015.05.020

Wang, Y. B., Wang, Y., Liu, X., and Tang, W. B. (2021). Research progress of photoperiod regulation in rice flowering. *Chinese J. Rice Sci.* 35, 207–224.

Watanabe, S., Hideshima, R., Xia, Z., Tsubokura, Y., Sato, S., Nakamoto, Y., et al. (2009). Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics* 182, 1251–1262. doi: 10.1534/genetics.108.098772

Watanabe, S., Xia, Z. J., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., et al. (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the GIGANTEA gene is involved in soybean maturity and flowering. *Genetics* 188, 395–407. doi: 10.1534/genetics.110.125062

Xia, Y. S., Ning, Z. S., Bai, C. H., Li, R. H., Yan, G. J., Siddique, K. H. M., et al. (2012). Allelic variations of a light harvesting chlorophyll a/b-binding protein gene (*Lhcb1*) associated with agronomic traits in barley. *PLoS ONE* 7, e37573. doi: 10.1371/journal.pone.0037573

Xia, Z. J., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., et al. (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus *E1* that regulates photoperiodic flowering. *Proc. Nati. Acad. Sci. U. S. A.* 109, E2155–2164. doi: 10.1073/pnas.1117982109

Xiao, D. L., Zhang, D., Ma, L., Wang, H. Y., and Lin, Y. Q. (2017). Preliminary study on differentially expressed genes of sparassis latifolia under light inducing. *Edible Fungi China* 36, 60–63. doi: 10.13629/j.cnki.53-1054.2017.05.014

Xu, Y. H., Liu, R., Yan, L., Liu, Q., Jiang, S. C., and Shen, Y. Y. (2012). Light-harvesting chlorophyll a/b-binding proteins are required for stomatal response to abscisic acid in *Arabidopsis*. *J. Exp. Bot.* 63, 1095–1106. doi: 10.1093/jxb/err315

Yamada, T., Teraishi, M., Hattori, K., and Ishimoto, M. (2001). Transformation of adzuki bean by agrobacterium tumefaciens. *Plant Cell Tiss. Org.* 64, 47–54. doi: 10.1023/A:1010635832468

Yang, W., Gong, R. G., Shi, J. J., Lai, J., Liao, M. A., and Liang, G. L. (2014). *De novo* assembly and functional annotation of the loquat young fruit transcriptome under chilling stress. *J Northwest Aand F U.* 42, 138–146. doi: 10.13207/j.cnki.jnwafu.2014.08.013

Yin, Z. C., Guo, W. Y., Liang, J., Xiao, H. Y., Hao, X. Y., Hou, A. F., et al. (2019). Effects of multiple N, P, and K fertilizer combinations on adzuki bean (*Vigna angularis*) yield in a semi-arid region of northeastern *China. Sci. Rep.* 9, 19408. doi: 10.1038/s41598-019-55997-9

Yun, T., Zhang, C., Hui, Y., Lu, X., Fang, J., Li, P., et al. (2011). Molecular Cloning and Characterization of a TRANSPORT INHIBITOR RESPONSE 1 (*TIR1*) from Nicotiana tabacum. *Russ. J. Plant Physiol.* 58, 149–156. doi: 10.1134/S1021443711010213

Zhai, H., Lv, S., Liang, S., Wu, H. Y., Zhang, X. Z., Liu, B. H., et al. (2014). GmFT4, a homolog of FLOWERING LOCUS T, is positively regulated by E1 and functions as a flowering repressor in soybean. *PLoS ONE* 9, e89030. doi: 10.1371/journal.pone.0089030

Zhang, L., Liu, Y., Han, L. L., Ma, Y. J., Ning, M., Xiao, K., et al. (2021). The effects of photoperiod-induction on development of apical inflorescence and anatomical property of floret in adzuki bean. *J. Agr. U. Hebei* 44, 21–27. doi: 10.13320/j.cnki.jauh.2021.0061

Zhang, L., Zhou, B., and Li, Y. H. (2010). Advances in the Structure and Function of HY5 in Plant. *Plant Physiol. Commun.* 46, 985–990. doi: 10.13592/j.cnki.ppj.2010.10.006

Zhang, X., Jiang, H., Wang, H., Cui, J., Wang, J. H., Hu, J., et al. (2017). Transcriptome analysis of rice seedling roots in response to potassium deficiency. *Sci. Rep.* 7, 5523. doi: 10.1038/s41598-017-05887-9

Zhang, Y. Y., Wang, Q., Li, X. X., Sun, T. T., Gong, D. P., Yang, M. L., et al. (2015). Analysis and fuctional prediction of MATE gene family incomm

on tobacco (*Nicotiana tabacum*). *J. Plant Genet. Resour.* 16, 1307–1314. doi: 10.13430/j.cnki.jpgr.2015.06.024

Zhao, C., Takeshima, R., Zhu, J. H., Xu, M. L., Sato, M., Watanabe, S., et al. (2016). A recessive allele for delayed flowering at the soybean maturity locus *E9* is a leaky allele of *FT2a*, a FLOWERING LOCUS T ortholog. *BMC Plant Biol.* 16, 20. doi: 10.1186/s12870-016-0704-9

Zhao, Y., Gao, X., Wang, D., Gao, C. P., and Yun, J. F. (2017). Cloning and expression analysis of Lhcb1 from *Agrophyron mongolicum* Keng under drought stress. *Acta Bot. Boreali-Occidentalia Sin.* 37, 211–216. doi: 10.7606/j.issn.1000-4025.2017.02.0211

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: 18 July 2022
doi: 10.3389/fpls.2022.946213

The Central Circadian Clock Protein TaCCA1 Regulates Seedling Growth and Spike Development in Wheat (*Triticum aestivum* L.)

Jie Gong[1,2†], Yimiao Tang[1,2†], Yongjie Liu[1,2], Renwei Sun[1,2], Yanhong Li[1,2], Jinxiu Ma[1,2], Shengquan Zhang[2], Fengting Zhang[2], Zhaobo Chen[2], Xiangzheng Liao[2], Hui Sun[2], Zefu Lu[3], Changping Zhao[1,2]* and Shiqing Gao[1,2]**

[1] The Municipal Key Laboratory of the Molecular Genetics of Hybrid Wheat, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [2] Institute of Hybrid Wheat, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China, [3] National Key Facility of Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China

OPEN ACCESS

Edited by:
Liang Wu,
Zhejiang University, China

Reviewed by:
Xiaoming Wang,
Northwest A&F University, China
Lu Liu,
Shanghai Jiao Tong University, China

***Correspondence:**
Zefu Lu
luzefu@caas.cn
Changping Zhao
changpingzhao@baafs.net.cn
Shiqing Gao
gshiq@126.com

† These authors have contributed equally to this work

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

Received: *17 May 2022*
Accepted: *20 June 2022*
Published: *18 July 2022*

Citation:
Gong J, Tang Y, Liu Y, Sun R, Li Y, Ma J, Zhang S, Zhang F, Chen Z, Liao X, Sun H, Lu Z, Zhao C and Gao S (2022) The Central Circadian Clock Protein TaCCA1 Regulates Seedling Growth and Spike Development in Wheat (Triticum aestivum L.). Front. Plant Sci. 13:946213. doi: 10.3389/fpls.2022.946213

The biological functions of the circadian clock on growth and development have been well elucidated in model plants, while its regulatory roles in crop species, especially the roles on yield-related traits, are poorly understood. In this study, we characterized the core clock gene *CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1)* homoeologs in wheat and studied their biological functions in seedling growth and spike development. *TaCCA1* homoeologs exhibit typical diurnal expression patterns, which are positively regulated by rhythmic histone modifications including histone H3 lysine 4 trimethylation (H3K4me3), histone H3 lysine 9 acetylation (H3K9Ac), and histone H3 lysine 36 trimethylation (H3K36me3). TaCCA1s are preferentially located in the nucleus and tend to form both homo- and heterodimers. *TaCCA1* overexpression (*TaCCA1*-OE) transgenic wheat plants show disrupted circadian rhythmicity coupling with reduced chlorophyll and starch content, as well as biomass at seedling stage, also decreased spike length, grain number per spike, and grain size at the ripening stage. Further studies using DNA affinity purification followed by deep sequencing [DNA affinity purification and sequencing (DAP-seq)] indicated that TaCCA1 preferentially binds to sequences similarly to "evening elements" (EE) motif in the wheat genome, particularly genes associated with photosynthesis, carbon utilization, and auxin homeostasis, and decreased transcriptional levels of these target genes are observed in *TaCCA1*-OE transgenic wheat plants. Collectively, our study provides novel insights into a circadian-mediated mechanism of gene regulation to coordinate photosynthetic and metabolic activities in wheat, which is important for optimal plant growth and crop yield formation.

Keywords: CCA1, EE motif, energy metabolism, histone modification, photosynthesis, wheat

INTRODUCTION

Photosynthesis is the basis of all life on Earth; it enables plants to take the sun's light energy and fix carbon from CO_2 to synthesize soluble sugars and starch providing energy for their growth and development (Stitt and Zeeman, 2012; Yamori and Shikanai, 2016). Recent molecular and genetic studies have identified several factors that regulate photosynthesis processes involved in

chlorophyll metabolism, photosystem assembly, and carbon fixation, all of which are crucial for plant growth and development. Pchlide oxidoreductases A and B (*PORA* and *PORB*) genes encode protochlorophyllide oxidoreductases a and b, which play roles in maintaining light-dependent chlorophyll biosynthesis in green plants (Reinbothe et al., 1996). Chlorophyll synthase (CHLG) acts in a salvage pathway for chlorophyll biosynthesis by re-esterifying the chlorophyllide a produced during chlorophyll turnover. A missense mutation in *CHLG* leads to a lower chlorophyll content and delayed chloroplast development in *Arabidopsis* and rice (Wu et al., 2007; Lin et al., 2014). Sucrose phosphate synthase (SPS) and sucrose phosphate phosphatase (SPP) are key enzymes involved in the synthesis of sucrose. Overexpressing *SPS* and *SPP* in *Arabidopsis* and hybrid poplar plants increase plant growth and biomass accumulation (Maloney et al., 2015). It has now been demonstrated that lots of genes, including *PORA*, *PORB*, *CHLG,* and *SPS,* are involved in circadian clock regulation in plants. In *Arabidopsis* and many important crops (e.g., rice and maize), biomass accumulation, photosynthesis, and seed number are increased by correct circadian regulation (Green et al., 2002; Dodd et al., 2005, 2014; Ni et al., 2009; Izawa et al., 2011; Haydon et al., 2013; Ko et al., 2016; Sanchez and Kay, 2016; Steed et al., 2021).

The circadian clock is an endogenous and autonomous timekeeping mechanism with a period of about 24 h that is found in most eukaryotic organisms. This clock allows organisms to perceive environmental cue changes (e.g., light and temperature) to help coordinate metabolic and developmental processes (Young and Kay, 2001; Harmer, 2009). In the model plant *Arabidopsis thaliana*, two partially redundant Myb-like transcription factors, LATE ELONGATED HYPOCOTYL (LHY) and CCA1, as well as the pseudo-response regulator (PRR) TIMING OF CAB EXPRESSION 1 (TOC1), represent central circadian clock components (Schaffer et al., 1998; Wang and Tobin, 1998; Strayer et al., 2000). When the endogenous clock matches the external diurnal cycle, photosynthetic activity, CO_2 fixation rates, and fitness are increased (Dodd et al., 2005).

CIRCADIAN CLOCK-ASSOCIATED 1 is a master clock protein that mediates rhythmic expression of the clock and output genes during plant growth and development. As examples, in *Arabidopsis*, *CCA1-OE* plants abolished expression rhythms and phases of clock output genes, leading to longer hypocotyls, delayed flowering time, lower chlorophyll content, reduced CO_2 assimilation, and reduced biomass compared to wild-type plants (Wang and Tobin, 1998; Dodd et al., 2005). In rice, OsCCA1 positively regulates expression of *TEOSINTE BRANCHED1 (OsTB1), DWARF14 (D14), and IDEAL PLANT ARCHITECTURE1 (IPA1)*, to repress tiller-bud outgrowth. Downregulating and overexpressing of *OsCCA1* increase and reduce tiller numbers, respectively. OsCCA1 also regulates *IPA1* expression to mediate panicle and grain development and could fine-tune photoperiodic flowering by directly regulating *OsGIGANTEA* (*OsGI*) (Wang et al., 2020; Sun et al., 2021). In maize and soybean, *ZmCCA1b* and four *GmLHYs* are homologous to *Arabidopsis CCA1/LHY*, playing important roles in chlorophyll metabolism, internode elongation, plant height, and drought stress responses (Ko et al., 2016; Cheng et al., 2019; Wang et al., 2021). In wheat, *TaLHY* is homologous to *Arabidopsis CCA1/LHY*, playing an important role in disease resistance against stripe rust fungus and ear heading (Zhang et al., 2015). However, the molecular basis for the clock genes to regulate wheat growth and development is largely unknown.

Wheat is one of most widely cultivated crops globally, and increasing wheat yield is known as a major pillar for producing sufficient food for a growing human population (Appels et al., 2018). So far, the molecular basis for the clock genes to regulate wheat growth and development is largely unknown. To explore the function and mechanism of action of the circadian clock in the wheat growth and development, we identified the wheat *TaCCA1* genes and characterized their roles in wheat growth and development. By generating and assessing wheat plants that overexpressing *TaCCA1*, functional assay results revealed key regulatory roles of TaCCA1 in wheat circadian rhythmicity, seedling growth, and spike development. We used DAP-seq to identify genome-wide binding sites of wheat TaCCA1 and correspondingly their direct downstream target genes. Multiple molecular biological assays showed that TaCCA1s target many output genes implicating in photosynthesis, energy metabolism, and auxin homeostasis through binding to the EE-like motif. Furthermore, TaCCA1 downregulated the expression of these output genes. To our knowledge, this investigation provides novel insights into a circadian-mediated mechanism for expressing genes in wheat to coordinate photosynthetic and metabolic activities, leading to optimal growth and development, then for the first time, establish the functions of wheat *CCA1* homoeologs and shed light to the circadian regulation in wheat.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The wheat cultivar 73064-1, a winter wheat and widely used in the wheat research community, and the model wheat cultivar "Fielder" were used in this study. All wheat plants were grown under short- (8 h light/16 h dark, 6:00 light on) and long- (16 h light/8 h dark, 6:00 light on) day conditions with a light intensity of 500 μmol/m^2/s and 25°C/20°C light/dark temperatures.

Growth Trait Measurement

For aerial biomass, above 10 ground seedlings for each line were collected at 10 days after planting and then weighed after desiccation for 48 h at 80°C. Spike length (SL) and kernel number per spike (KNS) are measured as the mean of the main spikes of 10 independent plants from each line. For seed weight, 100 kernels randomly selected from each line were weighed. For kernel length (KL) and kernel width (KW), 10 kernels randomly selected from main spike of each line were measured by manual. All experiments were replicated three times, unless noted otherwise.

Chlorophyll Content Measurement

Plant tissues' chlorophyll content was determined using a spectrophotometer according to a previously described method, with minor modification (Wu et al., 2007). Fresh leaves of WT and *TaCCA1-7D*-OE lines were cut and immersed in 10 ml ethanol for 48 h under dark conditions, following which plant

residuals were removed by centrifugation and the supernatants were analyzed with a UV-1800 Spectrophotometer (SHIMADZU, Japan) scanning at 470, 649, and 665 nm.

Starch Content Measurement

Leaf samples were weighed and immediately frozen in liquid nitrogen. The frozen tissues were ground, mixed with a homogenization buffer [500 mM 3-(N-Morpholino) propanesulfonic Acid (MOPS) pH 7.5, 5 mM Ethylenedinitrilotetraacetic acid (EDTA), and 10% ethyl glycol] and then filtered through Miracloth (CalBiochem, San Diego, CA, United States). After centrifugation, pellets were dissolved in DMSO to extract the insoluble carbohydrate fraction. The starch content was measured from the insoluble carbohydrate fraction using a commercial assay kit according to the manufacturer's instruction (R-Biopharm, Darmstadt, Germany).

RNA Extraction and qPCR

For the expression analysis of the target genes, total RNA was extracted from plant tissues using an RNAprep pure Plant Kit. The expression analysis of the target genes was performed using an Eco Real-Time PCR system (Illumina, San Diego, CA, United States). Wheat *Actin* gene (*TraesCS1B02G024500*) served as an internal control for the expression studies. The primers used are shown in **Supplementary Table 1**.

Sequence and Phylogenetic Analyses

The conserved domain of CCA1 proteins was predicted using the SMART software (Letunic and Bork, 2018). CCA1 protein sequences are HvCCA1 (HQ850270), BdCCA1 (LOC100838310), ZmCCA1a (ZM2G474769), ZmCCA1b (ZM2G014902), SbCCA1 (Sb07g003870), OsCCA1 (Os08g0157600), AtCCA1 (At2g46830), AtLHY (At1g01060), PnLHY1 (AB429410), PnLHY2 (AB429411), BraA.LHY.a (Bra030496), McCCA1 (AY371287), GmLCL1 (Glyma.16G017400), and GmLCL2 (Glyma.03G261800), Hv: *Hordeum vulgare*; Bd: *Brachypodium distachyon*; Zm, *Zea mays*; Sb: *Sorghum bicolor*; Os: *Oryza sativa*; At: *Arabidopsis thaliana*; Pn: *Populus nigra*; BraA: *Brassica rapa*; Mc: *Mesembryanthemum crystallinum;* and Gm: *Glycine max*. Full-length protein sequences were aligned using the Clustal W module (Larkin et al., 2007). The numbers of amino acid substitutions between each pair of LHY/CCA1 proteins were estimated by the Jones–Taylor–Thornton (JTT) model with the complete-deletion option (Jones et al., 1992). From estimated numbers of amino acid substitutions, the phylogenetic tree was reconstructed using the neighbor-joining method (Saitou and Nei, 1987). The bootstrap values were calculated with 1,000 replicates and shown next to the branches (Felsenstein, 1985). These analyses were performed using the MEGA7 software[1] (Kumar et al., 2016).

Chromatin Immunoprecipitation Assay Followed by qPCR

Chromatin immunoprecipitation (ChIP) assays were performed according to a previously described method (Li N. et al., 2018). Briefly, 2 g of samples were washed twice in cold phosphate-buffered saline (PBS) buffer, and proteins were cross-linked to DNA by incubating the samples with formaldehyde at a final concentration of 1% on a shaking device for 10 min at 4°C. Samples were then lysed, and chromatin was precipitated on ice. Chromatin was then sonicated to yield soluble sheared chromatin (200–500 bp). One part of the soluble chromatin was saved at −20°C for input DNA, and the remainder was used for immunoprecipitation with antibodies for H3K9Ac (CST, 9649), H3K4Me3 (CST, 9751), H3K36Me3 (CST, 4909), and normal rabbit IgG (CST, 2729). Immunoprecipitated DNA was amplified by PCR using their specific primers. PCR reactions were set up and run using the ChamQ SYBR Color qPCR Master Mix. The enrichment values were normalized to the input sample. The primers used are shown in **Supplementary Table 1**.

Subcellular Localization in Wheat Protoplasts

The coding sequence (CDS) (excluding the stop codon) of *TaCCA1s* was amplified with gene-specific primers and fused to the N-terminus of *GFP* in green fluorescent protein (GFP) expression vector (*CaMV35S-GFP-NOS*). Wheat protoplasts were isolated from the mesophyll tissue of wheat seedlings and then transformed using the polyethylene glycol (PEG) transfection method (Shan et al., 2014) separately with the plasmid DNA of *35S:TaCCA1-GFP*, and *35S:GFP* control as described previously. After PEG transfection, wheat protoplasts were incubated in W5 solution in a dark chamber at 23°C for 18 h, and GFP fluorescence was monitored under a laser-scanning confocal microscope. The primers used are shown in **Supplementary Table 1**.

Yeast Two-Hybrid Assays

The *TaCCA1-7A/B/D* CDSs were independently amplified and cloned into the destination vectors pGBKT7 or pGADT7. The yeast strain AH109 was transformed with these constructs using a lithium acetate transformation protocol (Yeast Protocols Handbook PT3024-1, Clontech). The positive colonies were selected SD/-Leu/-Trp (-LT) medium and then used for a growth assay on selective SD/-Leu/-Trp/-His/-Ade medium (-LTHA). The interactions were observed after 4 days of incubation at 30°C. The primers used are shown in **Supplementary Table 1**.

Biolayer Interferometry Assay

The binding affinity of TaCCA1-7A/7B for TaCCA1-7D and dsDNAs with TaCCA1-7D was measured using the ForteBio Octet RED 96e instrument. The TaCCA1-7A and TaCCA1-7B were diluted into 10–630 nM. Streptavidin (SA) biosensors were prewetted and loaded with 10.0 μg/ml biotinylated recombinant TaCCA1-7D. The equilibrium, association, and dissociation circles were carried out repetitively from low to high concentration of TaCCA1-7A/7B solutions. None of the protein loaded biosensors was used as the reference sensors for background subtraction. The sensors with the fixed TaCCA1-7D protein sample were applied to measure binding affinity of biotin-labeled double-strand DNA (dsDNA). Binding affinities

[1] http://www.megasoftware.net/index.html

were calculated as the ratio of dissociation and association rate constants using the Octet Data Analysis Software 11.0. The 5′-biotin-labeled oligonucleotides used are shown in **Supplementary Table 2**.

Generation of *TaCCA1*-OE Wheat Transgenic Lines

The overexpression construct was created by fusing the maize constitutive ubiquitin-1 (*ZmUBI-1*) promoter to the CDS of *TaCCA1-7D* in the binary vector *pLGY-02* (generously provided by Dr. Genying Li, Shandong Academy of Agricultural Science, Jinan, Shandong, China). The CDS of *TaCCA1-7D* was amplified with primers shown in **Supplementary Table 1**. The Plant Transformation Facility (Shandong Academy of Agricultural Science, Jinan, Shandong, China) transformed *pLGY-02-TaCCA1-7D* into the model wheat cultivar "Fielder" by *Agrobacterium*-mediated transformation according to their protocol.

Recombinant Protein Expression and Purification

The CDS of *TaCCA1-7A*/*7B*/*7D* was amplified and cloned into the pMAL-c2X vector (GE Healthcare). The primers used are shown in **Supplementary Table 1**. After the cell culture was induced with Isopropyl-β-D-thiogalactoside (IPTG), it was incubated at 18°C overnight. Bacteria were harvested and lysed with a high-pressure cell cracker. After centrifugation, the cleared extract was diluted and loaded on amylose-coupled agarose resin columns prepared according to the manufacturer's instruction (GE Healthcare). After columns were washed, maltose binding protein (MBP) and recombinant TaCCA1-7A/B/D were eluted and filtration by 0.22 μm filter unit (Millipore); the purified MBP and rTaCCA1-7A/B/D were aliquoted and stored at −80°C.

DAP-Seq Assay and Data Analysis

DAP-seq assay was performed with purified MBP fusion protein. Briefly, 10 μg MBP fusion proteins were incubated with 100 μl amylose-coupled agarose resin in 1 × PBS at 4°C for 1 h. The beads were collected by centrifugation at 500 *g* for 2 min. 1 μg sonicated DNA (300–1,000 bp) were then co-incubated with the beads in 1 × PBS + 0.1% NP40 at 25°C for 30 min. DNA was then purified, and libraries were constricted with Nextera DNA Library Prep Kit (Illumina FC-121-1031). Paired raw reads were trimmed with Trimmomatic-v.0.36 with default parameters. The remaining reads were aligned to ensemble 1.0 genome using Bowtie v2.3.4 with the following parameters: "bowtie2-X1000–very-sensitive." Aligned reads were sorted and filtered with MAPQ > 10 using SAMtools v1.9, and duplicated reads were removed using Picard version v2.16.0[2]. Model-based analysis of ChIP-seq 2 (MACS2) was used to define DNase I hypersensitive site (DHS) regions and DAP-seq peaks with the "keep-dup all" function. DAP-seq peaks were filtered with "Fold enrichment > 5 and FDR < 1e-5" and DHS peaks with "fold enrichment > 2 and FDR < 1e-2." DAP-seq peaks, which were overlapped >1 bp with DHS (bedtools v2.29.0), were further used. Peaks located within 5 kb upstream of transcriptional start sites (TSSs) were considered promoter-TSS; peaks located within 5 Kb downstream of transcriptional terminate sites (TESs) were consider as "TES"; genes overlapped with the gene body were considered "Genic"; and the rest ones were considered "Intergenic," and the closest genes were considered the associated genes. Gene Ontology (GO) enrichment analysis was performed using the EasyGO gene ontology enrichment analysis tool (Zhou and Su, 2007). The GO term enrichment was calculated using hypergeometric distribution with the *q*-value < 0.05. *Q*-values obtained by the Fisher's exact test were adjusted with a false discovery rate (FDR) for multiple comparisons to detect overrepresented GO terms.

[2]http://broadinstitute.github.io/picard/

Electrophoretic Mobility Shift Assay

Electrophoretic mobility shift assay was performed as previously described (Yang et al., 2019). The complementary oligonucleotides were annealed and labeled with DIG Gel Shift Kit (Roche). About 150 ng of MBP or TaCCA1-7D-MBP protein and 2 pM DIG-labeled probes were incubated in a 10 μl reaction mixture for 1 h on ice and then separated by 6% polyacrylamide gel in 0.5 × TBE buffer (pH 8.0) at 80 V for about 90 min. For competition assays, 10-, 20-, or 50-fold more non-labeled competitor DNA was mixed in the reaction before addition of the labeled probe. The oligonucleotides used are shown in **Supplementary Table 2**.

RESULTS

Identification of *TaCCA1* Genes in Bread Wheat

The CCA1 is a core component of the plant circadian clock and has been reported to play essential roles in regulating plant growth and development in *Arabidopsis*. To characterize the function of CCA1 ortholog in wheat, we identified three highly conserved homoeologous genes *TaCCA1-7A* (i.e., *TraesCS7A02G299400*), *TaCCA1-7B* (*TraesCS7B02G188000*), and *TaCCA1-7D* (*TraesCS7D02G295400*) in bread wheat genome. Further gene structure analysis revealed that the CDS of *TaCCA1-7A*, *TaCCA1-7B*, and *TaCCA1-7D* were 2,154, 2,136, and 2,157 bp in length, and they all contained 6 exons and 5 introns (**Supplementary Figure 1A**). The predicted TaCCA1-7A, TaCCA1-7B, and TaCCA1-7D proteins had 717, 711, and 718 amino acids, respectively, and all included a conserved SANT (SWI3, ADA2, N-CoR, and TFIIIB) DNA-binding domain at the N terminus (**Supplementary Figure 1B**). Phylogenetic analysis of CCA1 homologs and related sequences in angiosperms suggested high conservation of protein sequences in both monocot and eudicot species (**Supplementary Figure 2A**). Domain-level analysis also revealed that the SANT DNA-binding domain of the TaCCA1 proteins was more similarly to monocot SANT domains compared with eudicots (**Supplementary Figure 2B**).

TaCCA1 Genes Exhibit Diurnal Expression Patterns

To characterize the potentially functional conservation of *TaCCA1* genes in the circadian clockwork, we measured the expression level of three *TaCCA1* genes in wheat seedlings growing under both short- and long-day conditions. Seedlings were collected every 3 h from 6:00 a.m. (at dawn, Zeitgeber time 0, ZT0) over a 24-h period, and all three *TaCCA1* genes showed typical diurnal expression patterns under both short- and long-day conditions (**Figures 1A,B**). The transcripts of these genes began to increase gradually at midnight (ZT15) and reached peaks of diurnal rhythms around at dawn (ZT0) under short-day condition and early morning (ZT3) under long-day condition.

We next investigated the tissue-specific expression pattern of *TaCCA1* genes in wheat at ZT3 grown under long-day condition. All three *TaCCA1* genes exhibited similar expression patterns, with a strong accumulation in green tissues (**Figure 1C**), which were mostly the same as those observed in *LHY/CCA1* of *Arabidopsis* and *OsCCA1* of rice (Lu et al., 2009; Sun et al., 2021). Although *TaCCA1-7A*, *TaCCA1-7B*, and *TaCCA1-7D* all expressed in wheat seedlings at ZT3, *TaCCA1-7D* showed a relatively higher expression level than the other two homoeologs. Moreover, a similar expression profile for *TaCCA1s* was evident when we examined publicly available wheat expression profiling data (Ramirez-Gonzalez et al., 2018) (**Supplementary Figure 3**).

Collectively, these results support that the wheat *CCA1* genes exhibit conserved diurnal expression patterns and are expressed at significantly higher levels in the green tissues than the other tissues. As green tissues function in capturing energy from the sun and converting the energy into sugars and carbohydrates for plant growth and development, we speculated a potential role of *TaCCA1s* in wheat growth and development.

Diurnal Expression of *TaCCA1* Genes Is Associated With Histone Modification

Recent studies have revealed a link between circadian regulated gene expression and dynamic histone modifications. Core histones (i.e., H2A, H2B, H3, and H4) can be covalently modified at different positions with (among others) acetylation and methylation marks (Jenuwein and Allis, 2001). In *Arabidopsis*, *TOC1* expression is affected by clock-controlled cycles of histone acetylation (Perales and Mas, 2007), and both histone 3 acetylation (H3Ac) and H3K4me3 levels at the *LHY*, *CCA1*, and *TOC1* loci are positively correlated with the expression levels of their respective mRNA transcripts (Ni et al., 2009; Malapeira et al., 2012; Song and Noh, 2012; Chen and Mas, 2019). To examine the potential links between the rhythmic expression of the *TaCCA1* genes and dynamic histone modifications, we collected 14-day-old seedlings grown under long day (16 h light) conditions at ZT3, ZT9, and ZT15 and performed ChIP-qPCR and qPCR to detect the changes in histone modifications and gene expression levels, respectively. We examined chromatin changes in two different sites of the upstream regions (∼5,000 bp) of *TaCCA1* genes using antibodies against three marks for gene activation, namely, H3K4me3, H3K9ac, and H3K36me3 and then determined the expression levels of *TaCCA1* genes using qPCR. Examined together, the data revealed that the decreasing strength of *TaCCA1* genes expression observed from ZT3 to ZT9 and then to ZT15 was also reflected in decreased levels of H3K4me3, H3K9ac, and H3K36me3 marks at these three time points (**Figures 2A–D** and **Supplementary Figure 4**). These results indicate an apparent functional association between dynamic histone modifications with *TaCCA1* genes expression, indicating that *TaCCA1* gene expression is positively regulated by euchromatic histone marks H3K4me3, H3K9ac, and H3K36me3.

The TaCCA1 Proteins Are Mainly Localized in the Nucleus and Can Form Dimers

Recalling that the TaCCA1 proteins harbor the predicted SANT DNA binding domains, which mainly function in the nucleus, we examined their intracellular location by transiently expressing the *35S:TaCCA1s-GFP* fusion constructs in protoplasts isolated from wheat leaves. The fusion protein and a control expressing *35S:GFP* construct were monitored by

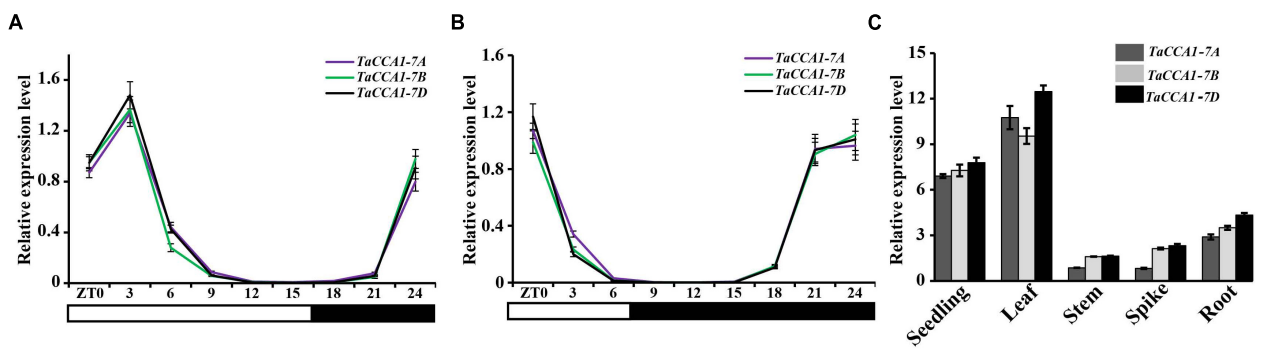


FIGURE 1 | Temporal and spatial expression of *TaCCA1* homoeologs. **(A,B)** qPCR analysis of relative expression levels of *TaCCA1-7A*, *TaCCA1-7B*, and *TaCCA1-7D* every 3 h in a 24-h period (light/dark cycle is shown below the histogram) in hexaploid wheat 73064-1 under short-**(A)** and long-day **(B)** conditions. **(C)** qPCR analysis of *TaCCA1* homoeologs relative expression levels in the seedling, leaf, stem, panicle, and root. Leaf, stem, panicle, and root are from booting stage wheat plants (line 73064-1) at ZT3. Values are means ± SD (n = 3).

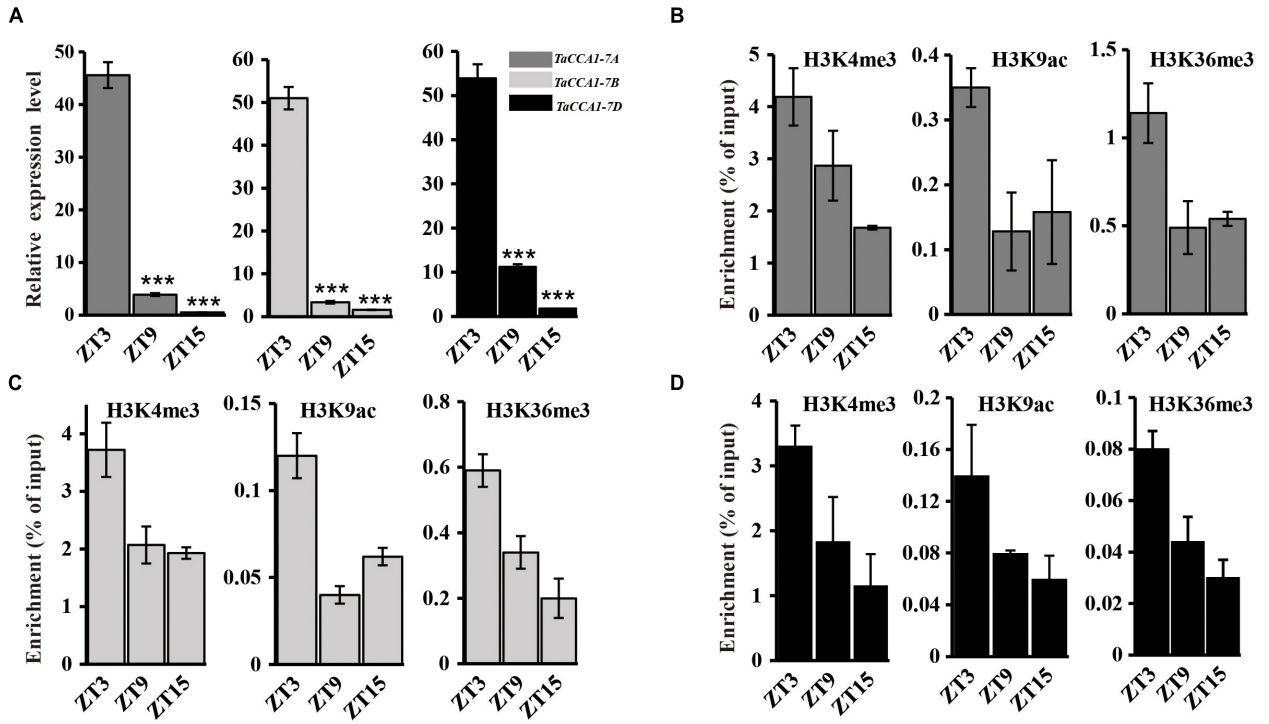


FIGURE 2 | Chromatin regulation of *TaCCA1* homoeologs. **(A)** qPCR analysis the expression of *TaCCA1* homoeologs at ZT3, ZT9, and ZT15. Values are means ± SD ($n = 3$). Asterisks indicate significant differences of gene expression between ZT3 and ZT9, ZT15 using Student's *t*-test (***$P < 0.001$). **(B–D)** ChIP-PCR analysis of *TaCCA1-7A*, *TaCCA1-7B,* and *TaCCA1-7D* promoter at ZT3, ZT9, and ZT15 using antibodies (Ab) against H3K4me3, H3K9Ac, and H3K36me3 in wheat line 73064-1. Values are means ± SD ($n = 2$).

confocal microscopy. Protoplasts transformed with *TaCCA1-7A-GFP* exhibited less fluorescence signals in the cytoplasm and cell membrane, majority of the fluorescence localized to the nucleus (**Figures 3A–H**). TaCCA1-7B and TaCCA1-7D showed the similar subcellular localization pattern with TaCCA1-7A (**Figures 3I–P**). These results suggest that the TaCCA1 proteins are mainly localized in the nucleus, supporting their putative function as transcription factors.

The *Arabidopsis* CCA1 can form dimers, which is functionally important for the maintenance of a roughly 24 h circadian rhythm (Lu et al., 2009; Yakir et al., 2009). We wonder whether TaCCA1 proteins take roles by forming homo- or heterodimers. Pursuing this hypothesis, we carried out yeast two-hybrid assays to characterize interactions among the TaCCA1 proteins and found that TaCCA1-7D can not only form homodimers independently but also form heterodimers with both TaCCA1-7A and TaCCA1-7B (**Figure 4A**). Moreover, the real-time binding graphs and kinetics parameters from the biolayer interferometry (BLI) assays further confirmed these interactions (**Figures 4B–D** and **Supplementary Table 3**). Together, these results indicate that besides the formation of homodimers as CCA1 does in *Arabidopsis*, wheat TaCCA1 proteins can also form heterodimers.

As the three *TaCCA1* homoeologs had high similarities in their expression pattern, amino acid sequences, and subcellular localization, we speculated the three *TaCCA1* homoeologs may have similar functions. Due to *TaCCA1-7D* exhibiting a relatively higher expression levels than *TaCCA1-7A* and *TaCCA1-7B* in wheat green tissues, as well as TaCCA1-7D showing closer evolutionary relationship with OsCCA1 and ZmCCA1s than TaCCA1-7A and TaCCA1-7B (**Supplementary Figure 2A**), it was selected for further functional analysis.

Overexpressing *TaCCA1-7D* Represses Seedling Growth and Reduces Yield in Wheat

To explore the biological functions of *TaCCA1s* in wheat, we generated *TaCCA1-7D* overexpression (*TaCCA1-7D*-OE) plants by driving *TaCCA1-7D* under the control of a constitutive *ZmUBI-1* promoter in wheat (cv. Fielder) and obtained 14 independent transgenic lines. Three *TaCCA1-7D*-OE lines, in which the *TaCCA1-7D* expression levels were obviously higher than the WT Fielder, were selected for further analysis (**Figure 5B**). In greenhouse conditions, compared with WT, all the three lines of *TaCCA1-7D*-OE seedlings exhibited shorter statures and a 19–37% decrease in biomass (**Figures 5A,E**). Also, physiological assay results indicated that transgenic lines had lower chlorophyll and starch contents than WT, all playing important roles in growth and development in plant (**Figures 5C,D**). Furthermore, *TaCCA1-7D*-OE plants had reduced SLs (decreased by 20–28%) and KNS (decreased

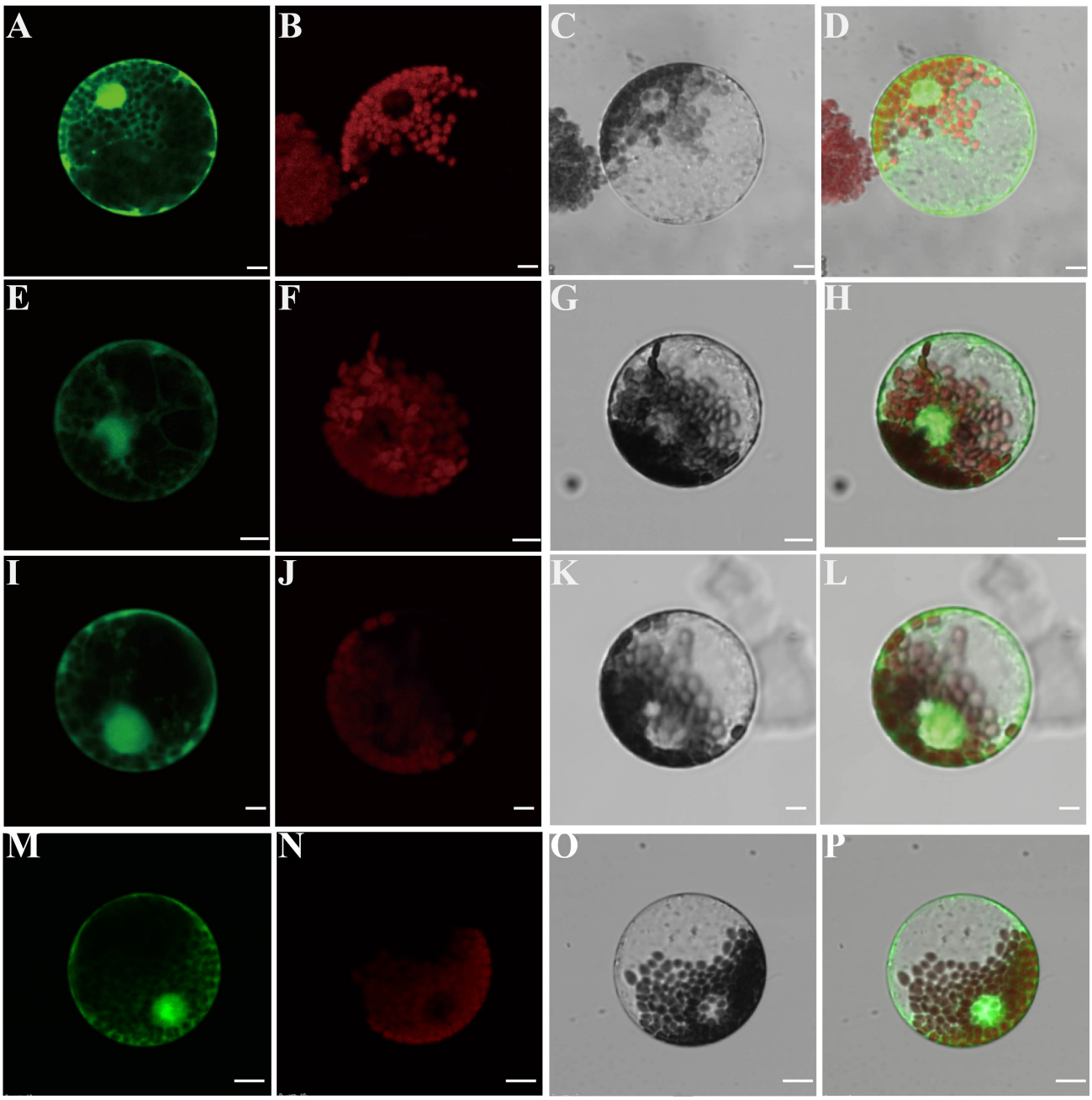


FIGURE 3 | The TaCCA1 proteins are mainly localized in the nucleus. **(A–D)** Wheat protoplasts expressing green fluorescence protein (GFP) alone. **(A)** Green fluorescent protein (GFP) fluorescence. **(B)** Autofluorescence of chloroplasts. **(C)** Bright-field image of GFP. **(D)** Merged image of **(A–C)**. **(E–H)** Wheat protoplasts cell expressing the TaCCA1-7A-GFP fusion protein. **(E)** GFP fluorescence. **(F)** Autofluorescence of chloroplasts. **(G)** Bright-field image of GFP. **(H)** Merged image of **(E–G)**. **(I–L)** Wheat protoplasts cell expressing the TaCCA1-7B-GFP fusion protein. **(I)** GFP fluorescence. **(J)** Autofluorescence of chloroplasts. **(K)** Bright-field image of GFP. **(L)** Merged image of **(I–K)**. **(M–P)** Wheat protoplasts cell expressing the TaCCA1-7D–GFP fusion protein. **(M)** GFP fluorescence. **(N)** Autofluorescence of chloroplasts. **(O)** Bright-field image of GFP. **(P)** Merged image of **(M–O)**. Scale bars = 10 μm.

by 30–40%) compared with WT plants at the ripening stage (**Figures 5F–H**). In addition, *TaCCA1-7D-OE* plants also had reduced tiller number (decreased by 55%), KLs (decreased by 7–10%), KWs (decreased by 11–13%), and 100-seed weight (decreased by 19–33%) compared with WTs, which were also important yield traits (**Figure 5I** and **Supplementary Figures 5A–D**). Consistent with the higher *TaCCA1-7D* expression level, we found that the *TaCCA1-7D-OE-1* plants showed more severe phenotypes than the other two lines.

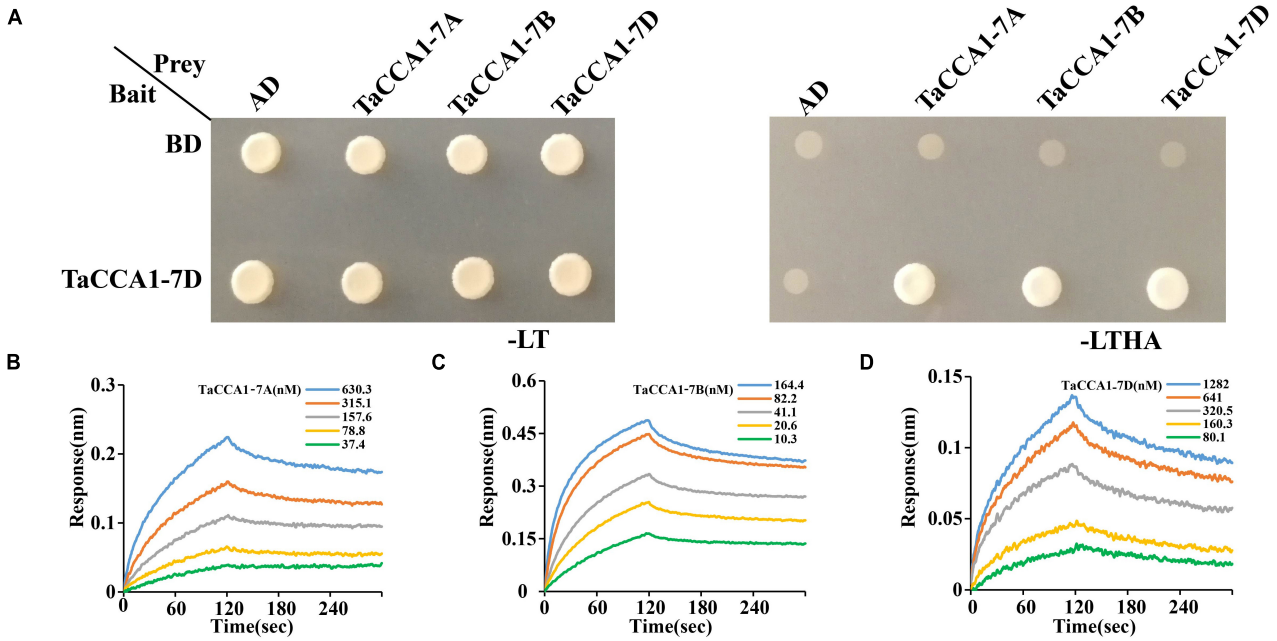


FIGURE 4 | TaCCA1s can form homodimers and heterodimers *in vitro*. **(A)** Interaction analysis between TaCCA1-7D and TaCCA1-7A, as well as TaCCA1-7B and TaCCA1-7D in yeast cells. Yeast cells growing on selective media without Leu, Trp, His, and Ade (-LTHA) represent positive interactions. **(B–D)** Biolayer interferometry confirms the interaction between TaCCA1-7D with TaCCA1-7A, TaCCA1-7B, and TaCCA1-7A.

Overexpressing *TaCCA1-7D* Disrupts Circadian Rhythmicity in Wheat

As known, CCA1 is an essential component of the circadian clock that functions to maintain circadian rhythmicity in *Arabidopsis* (Wang and Tobin, 1998). To determine the effects of *TaCCA1* overexpression on circadian rhythmicity in wheat, we measured the 24-h period time course expression of circadian clock genes including *TaPRR73* and *TaLUX ARRHYTHMO* (*TaLUX*) in WT, and *TaCCA1-7D*-OE-1 plants under long-day condition. Indeed, the rhythmic expression of these two clock genes was disrupted by constitutive expressing *TaCCA1-7D* (**Figures 5J,K**). Moreover, we checked the expression of *TaCCA1* homoeologs and other core clock genes in *TaCCA1-7D*-OE-1 plant at ZT3 under long-day condition. We found that the expression level of *TaCCA1-7A*, *TaCCA1-7B,* and endogenous *TaCCA1-7D* all reduced in these plants, indicating the potential feedback regulations of *TaCCA1* homoeologs (**Supplementary Figure 6**). Reduced expression of *TaTOC1* and elevated expression of *TaPRR37*, *TaPRR73,* and *TaPRR95* were also found in *TaCCA1-7D*-OE-1 plants (**Supplementary Figure 6**). Together, these results establish that *TaCCA1-7D* functions are essential in the maintenance of circadian rhythmicity in wheat.

TaCCA1-7D Preferentially Regulates Some Downstream Denes Involved in Photosynthesis and Carbon Utilization by Binding EE-Like Motifs

To explore the underlying regulation networks of TaCCA1s, the typical circadian rhythm-control transcription factor in wheat, we performed genome-wide TaCCA1-7D-binding profiles using DNA affinity purification followed by DAP-seq. Two independent DAP-seq experiments were highly correlated and showed highly Fraction of Reads in Peaks (FRiP) values (**Supplementary Figure 7**). The mapping statistics of DAP-seq data and distribution of DAP-seq peaks are shown in **Supplementary Tables 4, 5**. Transcription factor binding events usually happen in the accessible chromatin regions. Therefore, we further identified 7,284 DAP-seq peaks located in the DHSs from a previous study (Li et al., 2019) (**Figure 6A**). Totally, 69% TaCCA1-binding sites were located in the intergenic regions, 18% in promoter-TSS, 8% in TES, and 5% in gene bodies (**Figure 6B**). *De novo* motif analysis found a top-scoring motif (DRATATCH) that is similar in sequence to the well-known, clock-associated, "evening element (EE)" motif (AAATATCT) (**Figure 6C**). Both the total DAP-seq peaks and the DHS-filtered TaCCA1-binding sites enriched AATATC motifs (**Figure 6D**), suggesting the high quality of DAP-seq peaks. We further checked the profiling of TaCCA1-binding sites from 2-kb upstream of TSS to 2-kb downstream of TES, which showed TaCCA1 binding sites were mostly enriched in the 1-kb upstream of TSS and 1-kb downstream of TES (**Figures 6E,F**).

Furthermore, we used the GO enrichment analysis tool catalog DAP-seq peaks to enable exploratory analysis about the potential biological impacts of TaCCA1-mediated transcriptional regulation. GO analysis indicated a strong enrichment for genes with predicted functions relating to carbon utilization, photosynthesis, and metabolic process (**Figure 7A**). Consistent with the relatively high levels of *TaCCA1* expression in green tissues (**Figure 1C**), this result emphasizes that the transcriptional

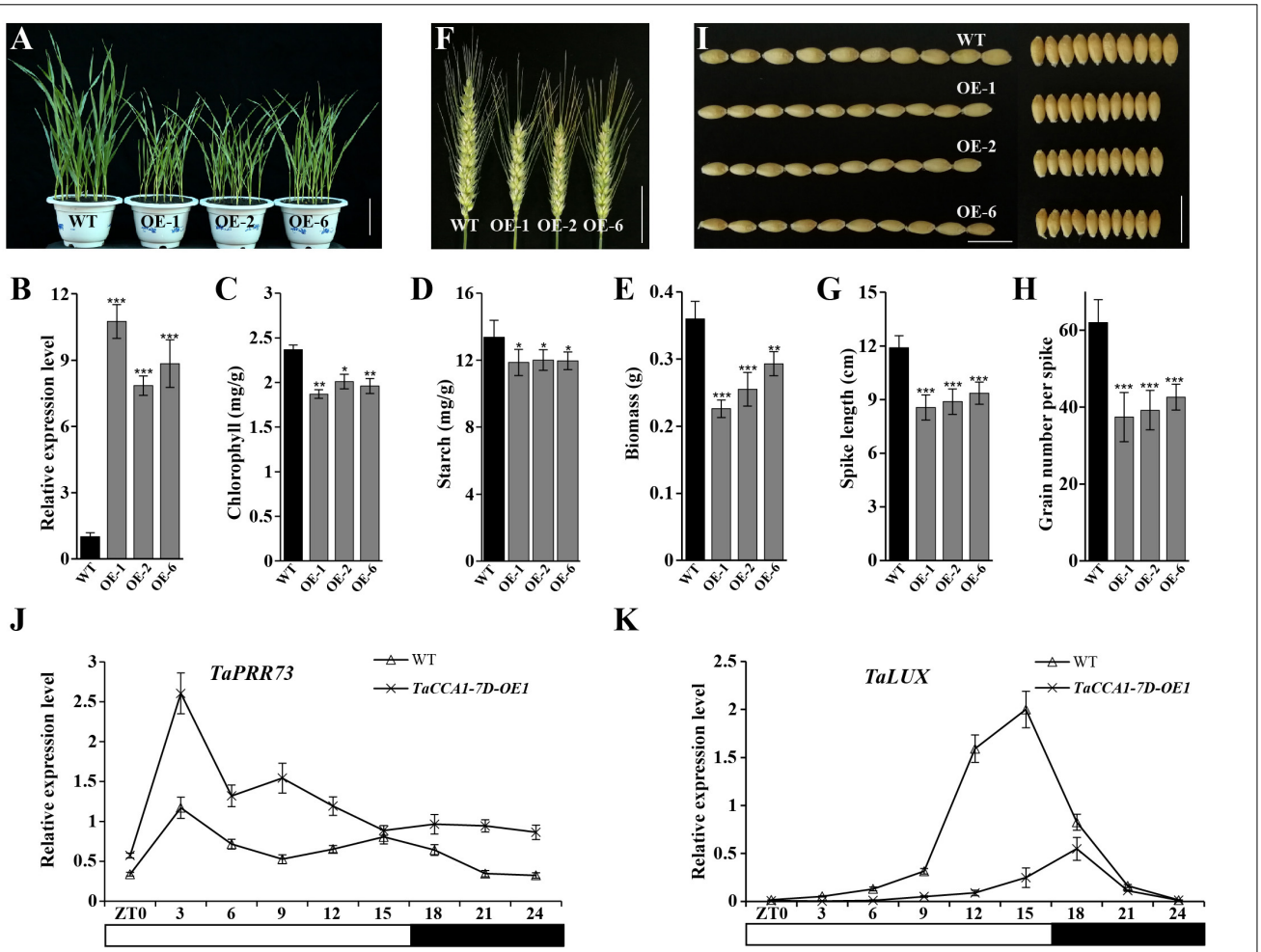


FIGURE 5 | Phenotypes of *TaCCA1-7D*-OE transgenic wheat. **(A)** Representative phenotype of wild-type (WT) (Fielder) and *TaCCA1-7D*-OE plants at the seedling stage. Scale bar = 5 cm. **(B)** qPCR analysis of *TaCCA1-7D* expression in 14-day-old WT and *TaCCA1-7D*-OE plants. Values are means ± SD (n = 3). **(C–E)** Reduced chlorophyll accumulation, starch accumulation, and dry aerial biomass of 14-day-old WT and *TaCCA1-7D*-OE-1, OE-2, and OE-6 plants. Values are means ± SD (n = 10). **(F)** Representative spike phenotypes of WT and *TaCCA1-7D*-OE plants at the ripening stage. Scale bar = 5 cm. **(G)** Spike length of WT and *TaCCA1-7D*-OE plants. Values are means ± SD (n = 10). **(H)** Grain number per spike of WT and *TaCCA1-7D*-OE plants. Values are means ± SD (n = 10). **(I)** Representative grain length (left) and width (right) phenotypes of WT and *TaCCA1-7D*-OE plants at the ripening stage. Scale bar = 1 cm. **(J,K)** qPCR analysis the expression of circadian genes *TaPRR73* and *TaLUX* in WT and *TaCCA1-7D*-OE1 plants. Asterisks indicate significant differences between transgenic plants and WT at the same development stages using Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001).

regulation activity of TaCCA1-7D, and potentially the other TaCCA1 proteins, might play an important role in photosynthesis and energy metabolism.

Finally, BLI and EMSA assays were used to further confirm the binding of TaCCA1-7D at promoters of some candidate genes involved in these pathways. Specifically, we tested binding at the genes encoding CHLG (*TraesCS1D02G226100*) involved in chlorophyll synthesis, component protein of Photosystem II (*TraesCS3B02G186700*) and Photosystem II reaction center protein L (*TraesCS1D02G295200*) involved in photosynthesis, and starch synthase (*TraesCS2B02G491700*) involved in starch synthesis. EMSA assays have confirmed the binding activities of TaCCA1-7D to the promoters of *TraesCS2B02G491700* and *TraesCS1D02G295200* (**Figure 7B**). BLI assays also have confirmed the binding activities of TaCCA1-7D to the promoters of *TraesCS1D02G226100*, *TraesCS3B02G186700*, *TraesCS1D02G295200,* and *TraesCS2B02G491700*. TaCCA1-7D showed similar binding pattern to the promoter of these genes with single digit nano molar affinity, and the real-time binding graphs and kinetics parameters are shown in **Supplementary Figure 8** and **Supplementary Table 6**.

To explore the underlying molecular mechanisms of *TaCCA1-7D-OE* phenotypes, we measured the expression of *TraesCS1D02G226100*, *TraesCS3B02G186700*, *TraesCS1D02G295200,* and *TraesCS2B02G491700* in the WT and *TaCCA1-7D-OE* lines. The results showed that the four genes all exhibited lower expression in *TaCCA1-7D-OE* lines than in WT at ZT6, consistent with the lower chlorophyll and starch contents in *TaCCA1-7D-OE* lines (**Figure 7C**). All these results support the notion that *TaCCA1s* take significant roles

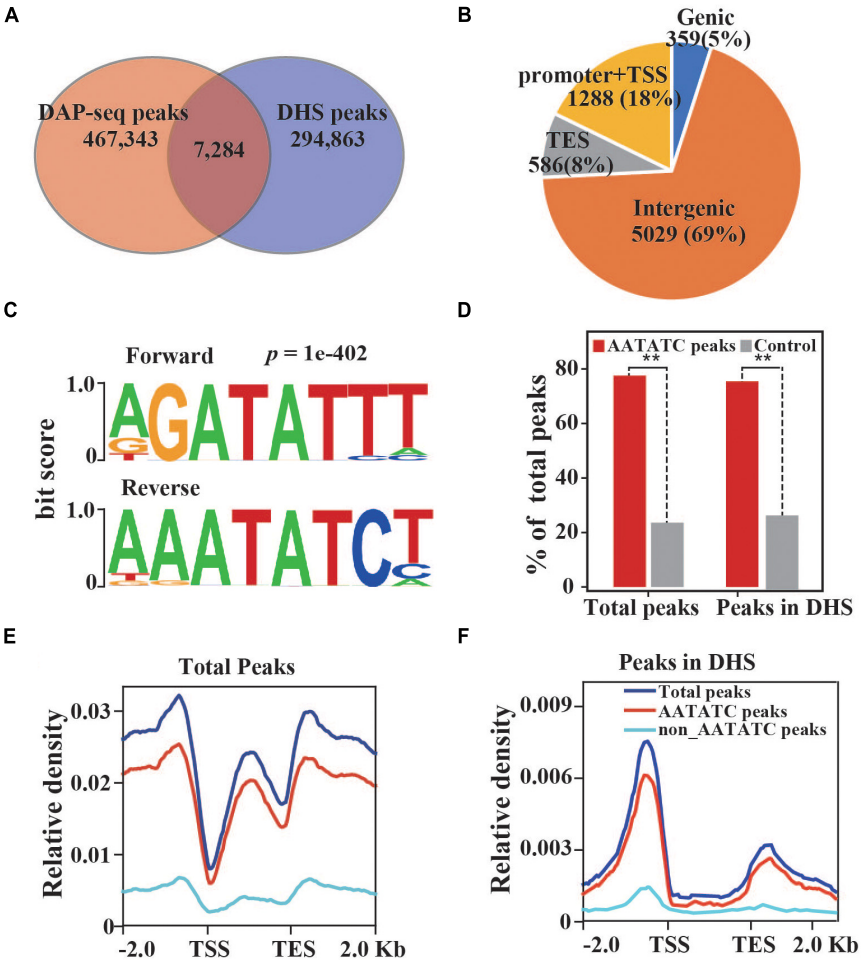


FIGURE 6 | Identification of TaCCA1-7D binding loci with DNA affinity purification and sequencing (DAP-seq). **(A)** An overlap between DAP-seq peaks and DNase-I hypersensitive sites (DHSs). The overlapped peaks were used for the following analysis. **(B)** Distribution of filtered DAP-seq peaks in different genomic regions. **(C)** Enriched DNA binding elements in DAP-seq peaks. **(D)** AATATC motifs were significantly enriched in DAP-seq peaks. The shuffled genomic regions were used. **(E,F)** Distribution of DAP-seq peaks **(E)** and filtered DAP-seq peaks **(F)** in ±2 kb of genes. Asterisks indicate significant difference of enriched AATATC peaks between TaCCA1 DAP-seq and control by the Fisher's exact test (**$P < 0.01$).

in regulating wheat seedling growth and yield-related traits through mediating transcriptional repression of a suite of genes involved in photosynthesis and energy metabolism pathways in plants. Thus, constitutive expressing *TaCCA1* in *TaCCA1-7D-OE* transgenic wheat led to repressed seedling growth and decreased yield.

DISCUSSION

Circadian clock regulating growth and development in *Arabidopsis* has been comprehensively surveyed, while comparatively less is known in crop species. In this study, we conducted an initial characterization of the wheat core circadian clock gene *TaCCA1s* and explored their regulation mechanisms and contributions to seedling growth and spike development in wheat, one of the most widely cultivated cereals globally. Our work establishes that the wheat TaCCA1 proteins all mainly localize to the nucleus and can dimerize into homo- and heterodimers. It has been shown that heterodimerization of central clock components contributes to the regulation of clock function in many circadian systems (Dunlap, 1999). Examples include the interaction of Period (PER) and Timeless (TIM) in *Drosophila* as a means of controlling the nuclear entry and degradation of both proteins (Edery, 1999; Meyer et al., 2006), heterodimers of the transcription factors mouse circadian locomotor output cycles kaput (mCLOCK), and brain and muscle Arnt-like protein-1 (BMAL1) in mammals accelerating their turnover, as well as E-box-dependent clock gene transcription (Reppert and Weaver, 2002; Lowrey and Takahashi, 2004) and interactions between CCA1 and LHY in *Arabidopsis* function synergistically in regulating circadian rhythms (Lu et al., 2009; Yakir et al., 2009). In this study, we demonstrated that TaCCA1 proteins all mainly localize in the nucleus and can physically interact with each other

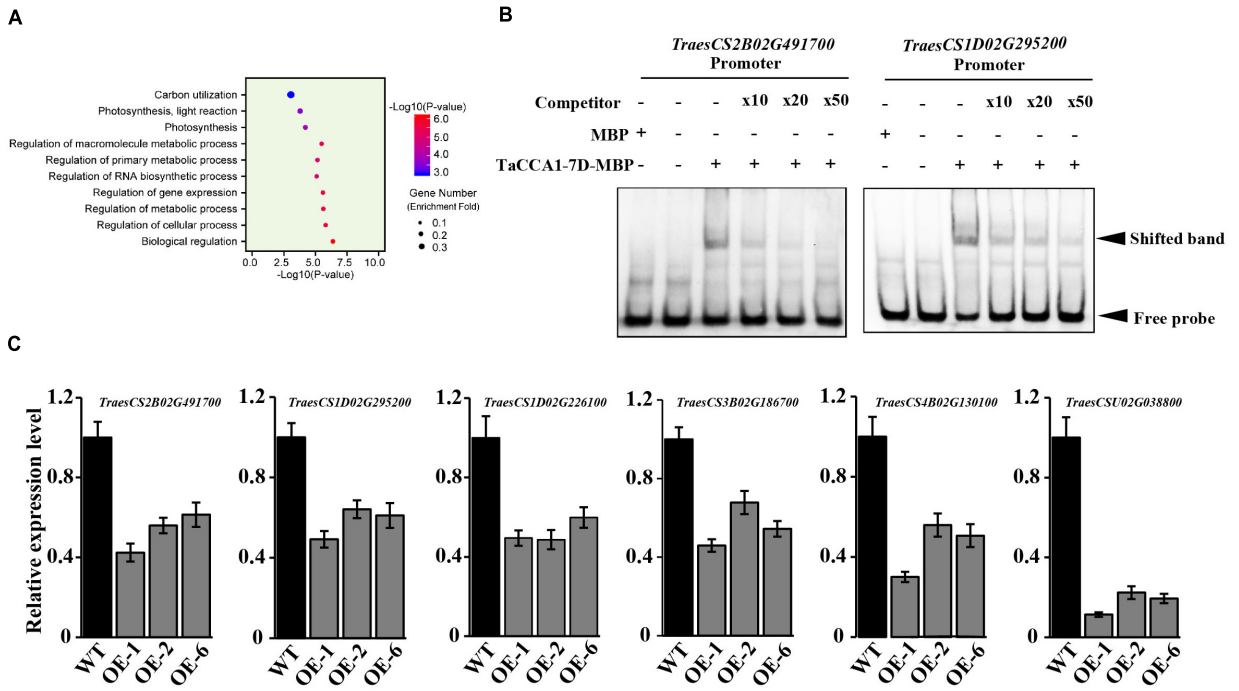


FIGURE 7 | Wheat TaCCA1 regulates genes involved in photosynthesis, energy metabolism, and auxin synthesis and transport *via* binding to EE-like motifs. **(A)** Functional enrichment analysis against GO Slim of the 15,675 putative CCA1 targets identified with a peak within 5 kb of the promoter and transcriptional start site (TSS). **(B)** Electrophoretic mobility shift assay (EMSA) analysis of TaCCA1-7D binding to the EE-like motif in *TraesCS2B02G491700* and *TraesCS1D02G295200* promoter. Lane 1, labeled probe incubated with MBP; lane 2, only labeled probe loaded; lane 3, labeled probe incubated with recombinant MBP-TaCCA1-7D; lanes 4 to 6, excessive unlabeled probe was added as a competitor with the following competitor: probe ratios: 10 (lane 4), 20 (lane 5), and 50 (lane 6). **(C)** qPCR analysis of relative expression levels (means ± SEM, $n = 3$) of EE-motif-containing genes *TraesCS2B02G491700, TraesCS1D02G295200, TraesCS1D02G226100, TraesCS5D02G464800, TraesCS4B02G130100,* and *TraesCSU02G038800* in wild-type (Fielder), and *TaCCA1-7D-OE* lines grown under LD at ZT6.

forming homo- or heterodimers, suggesting that dimerization is apparently conserved aspects in the regulation of eukaryotic circadian clocks. Nevertheless, the underlying mechanisms of dimer formation *in vivo* and functional divergence between homo- and heterodimerization of TaCCA1 proteins in wheat are still needed in future studies.

TaCCA1-OE transgenic wheat plants show disrupted circadian rhythmicity coupling with reduced chlorophyll and starch contents, as well as biomass at the seedling stage. Meanwhile, decreased SL and grain number per spike at the ripening stage. Consistently, in allotetraploid hybrids of the eudicot *Arabidopsis* and intraspecific hybrids of the monocot maize, altered expression of *CCA1* has been implicated in increased photosynthetic and metabolic activities. In *A. thaliana*, overexpression of *CCA1* causes disrupted circadian rhythms and reduces both photosynthesis activity and fitness (Green et al., 2002). Double-mutant *cca1 lhy* plants accumulate less starch and are unable to properly regulate the rate of starch degradation to match the length of night (Graf et al., 2010). In maize, overexpressing *ZmCCA1b* reduced chlorophyll content and plant height (Ko et al., 2016). In rice, downregulating and overexpressing *OsCCA1* increases and reduces tiller numbers, respectively. *OsCCA1* also mediates panicle and grain development and could fine-tune photoperiodic flowering. Gene expression analysis involved in photosynthesis and energy metabolism pathways in *TaCCA1-OE* transgenic implied the important regulation roles of *TaCCA1* genes. Prior to our study, it was not clear whether similar mechanism(s) may mediate the growth and development of wheat.

Mechanistically, we identified genome-wide direct targets of TaCCA1 and its binding core elements (EE-like motif) in wheat for the first time, which will be a valuable resource to systematically uncover TaCCA1s-mediated circadian outputs. We identified that TaCCA1s preferentially bound at promoters harboring the EE-like motif in wheat and confirmed the binding activities using BLI and EMSA assays. Downregulation of photosynthesis- and energy metabolism-related genes in *TaCCA1-OE* lines implies the important regulation roles of *TaCCA1* homoeologs in wheat seedling growth and spike development. Clock-related improvements in fitness results from appropriate temporal regulation of energetically costly activities, including phytohormone synthesis and signaling, growth control, metabolic activities, plant–pathogen interactions, and abiotic stress responses (Covington and Harmer, 2007; Hotta et al., 2007; Nozue et al., 2007; Covington et al., 2008; Gutierrez et al., 2008; Legnaioli et al., 2009; Wang et al., 2011; Goodspeed et al., 2012). The regulation of each of these processes is disrupted in the circadian arrhythmic *A. thaliana* lines overexpressing *CCA1*. Beyond demonstrating circadian impacts on seedling growth and spike development

in wheat, our study establishes functional conservation among the *Arabidopsis*, maize, rice, and wheat CCA1 orthologs. Consistent with the regulatory role of CCA1 in *Arabidopsis* and maize, we showed that TaCCA1 can bind to the promotors of photosynthetic and metabolic activity related genes through EE-like motif. Further exploration of the functions of these EE-motif-containing CCA1 target genes will reveal the mechanisms through which circadian clocks regulate unique and complex biological pathways fine-tuning the growth and development in wheat.

In addition to the conserved mechanisms in *Arabidopsis* and rice, key genes involved in auxin homeostasis and transport were identified as a potential target of TaCCA1s. We found that the promoters of *TraesCS4B02G130100* (encoding PIN, auxin efflux carrier) and *TraesCSU02G038800* (encoding YUCCA, indole-3-pyruvate monooxygenase) harbor EE-like motifs and had apparent enrichment in our DAP-seq experiment. Furthermore, we measured the expression levels of these two genes in the WT Fielder and *TaCCA1-7D-OE* (i.e., OE-1, OE-2, and OE-6) lines. The results showed that the two genes exhibited

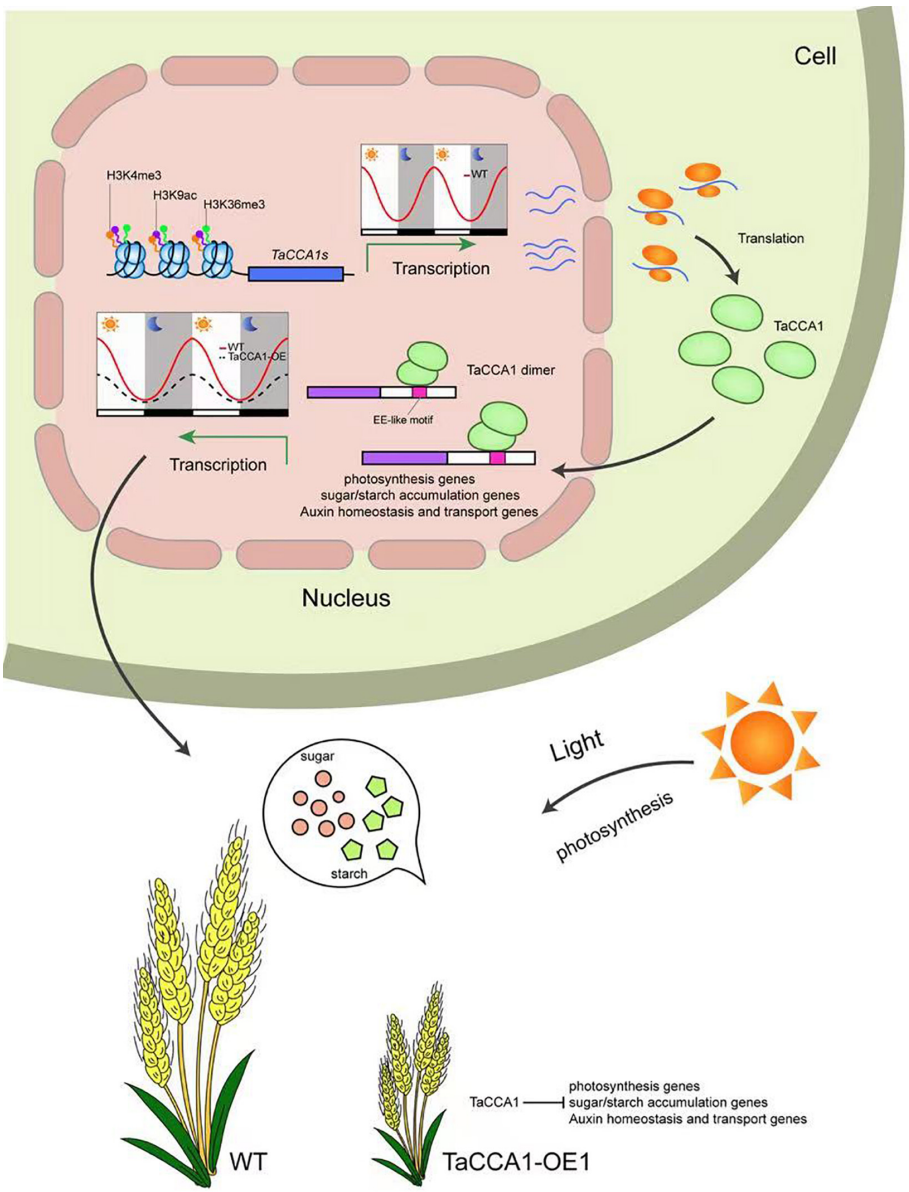


FIGURE 8 | Working model of possible molecular mechanism of the central circadian clock protein TaCCA1-mediated pathways involving in seedling growth and spike-related traits in wheat. *TaCCA1* homoeologs exhibit typical diurnal expression patterns, which are positively regulated by rhythmic histone modifications. TaCCA1s are preferentially located in the nucleus and tend to form homo- or heterodimers. Overexpressing of *TaCCA1* downregulates genes implicating in photosynthesis, energy metabolism, and auxin homeostasis through binding EE-like motifs, restricting the seedling growth and development of spike-related traits.

lower expression in *TaCCA1-7D-OE* lines than in WT Fielder (**Figure 7C**). Phytohormones, including indole acetic acid (IAA and auxin), are key regulators in promoting spike development. IAA plays essential roles in axillary meristem (AM) initiation and outgrowth by promoting cell polarity establishment and cell elongation, resulting in a distinct plant architecture, panicle type, and grain number per spike (Yue et al., 2017; Li Y. et al., 2018). IAA is also a factor stimulating assimilate transport to developing grain (Darussalam et al., 1998). Photoassimilate and dry matter import into developing grains of wheat by affecting components of grain growth, such as cell enlargement and nutrient accumulation, are enhanced by increasing IAA levels (Hess et al., 2002). Moreover, IAA plays important roles in other plant growth stages. In *Catharanthus roseus*, IBA, IAA, and NAA significantly improved plant fresh and dry weights, total chlorophyll and carotenoids content, and net photosynthetic rate (Masidur Alam et al., 2012). In *Panax ginseng*, exogenous IAA enhanced significantly the net photosynthetic rate, stomatal conductance, and transpiration rate at all growth stages (Li and Xu, 2014). Recalling the phenotypes (especially spike-related traits) of *TaCCA1-7D-OE* lines, we thought that genes involved in wheat IAA homeostasis and transport regulated by TaCCA1 played important roles in seedling growth and spike development coupling with photosynthesis and energy metabolism pathways.

Over the past decade, the seminal findings in model plant *Arabidopsis* have been used to link circadian clock gene function to agronomic traits across many eudicots and monocots (Bendix et al., 2015). Circadian genes from crops have been reported to modulate many key agriculture traits, such as tillering and flowering in rice, heterosis in maize, nodulation in legumes, tuberization in potato, postharvest management of fruit, and vegetable storage (Yanovsky et al., 2000; Goodspeed et al., 2013; Ko et al., 2016; Kong et al., 2020; Wang et al., 2020; Sun et al., 2021). Nonetheless, there remain critical knowledge gaps related to the molecular components of circadian rhythms in important crop groups and fail to widely apply circadian studies to agricultural production (Steed et al., 2021). Based on this, what urgently needs to be solved in the future is there is a need for fundamental research in the crops to understand how the circadian clock helps crops to anticipate changes in extreme condition they encounter, as well as how innovative plant breeding may improve to produce beneficial agronomic traits.

CONCLUSION

In wheat, the central clock proteins TaCCA1s target many output genes implicating in photosynthesis, energy metabolism and auxin homeostasis through binding EE-like motif. TaCCA1s regulate the expression of genes involved in carbon assimilation and energy metabolism, which is established early in the seedling and subsequently maintained during growth. Our findings thus provide novel insights into a circadian-mediated mechanism for expressing genes in wheat to coordinate photosynthetic and metabolic activities, leading to optimal growth and development (**Figure 8**). In summary, our findings extend our knowledge of plant circadian clock and provide evidence for their functional roles in seedling growth and spike-related traits in the important staple crop wheat.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are publicly available. This data can be found here: NCBI, GSE200426.

AUTHOR CONTRIBUTIONS

JG conceived the project. JG, SG, and CZ designed the experiments. JG performed most of the experiments with help from YT, RS, YjL, YhL, JM, SZ, FZ, ZC, XL, and HS. ZL performed the DAP-seq experiment and data analysis. JG and SG analyzed the data and wrote the article. All authors approved the article for publication.

FUNDING

This work was supported by the Natural Science Foundation of Beijing (grant no. 6194034), the Important Crops Genetically Modified New Germplasm Creation of BAAFS (grant no. KJCX20200205), and Collaborative Innovation Center for Crop Phenomics (grant no. KJCX201917).

ACKNOWLEDGMENTS

We sincerely thank Genying Li for helping us to establish wheat transformation system. We also thank Ming Chen from Chinese Academy of Agricultural Sciences and Rongxin Yang from Nanchang University for critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.946213/full#supplementary-material

REFERENCES

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Bendix, C., Marshall, C. M., and Harmon, F. G. (2015). Circadian clock genes universally control key agricultural traits. *Mol. Plant* 8, 1135–1152. doi: 10.1016/j.molp.2015.03.003

Chen, Z. J., and Mas, P. (2019). Interactive roles of chromatin regulation and circadian clock function in plants. *Genome Biol.* 20:62. doi: 10.1186/s13059-019-1672-9

Cheng, Q., Dong, L., Su, T., Li, T., Gan, Z., Nan, H., et al. (2019). CRISPR/Cas9-mediated targeted mutagenesis of *GmLHY* genes alters plant height and internode length in soybean. *BMC Plant Biol.* 19:562. doi: 10.1186/s12870-019-2145-8

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

Covington, M. F., Maloof, J. N., Straume, M., Kay, S. A., and Harmer, S. L. (2008). Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biol.* 9:R130. doi: 10.1186/gb-2008-9-8-r130

Darussalam, M., Cole, M. A., and Patrick, J. W. (1998). Auxin control of photoassimilate transport to and within developing grains of wheat.pdf. *Aust. J. Plant Physiol.* 25, 69–77. doi: 10.1071/PP97080

Dodd, A. N., Kusakina, J., Hall, A., Gould, P. D., and Hanaoka, M. (2014). The circadian regulation of photosynthesis. *Photosynth. Res.* 119, 181–190. doi: 10.1007/s11120-013-9811-8

Dodd, A. N., Salathia, N., Hall, A., Kevei, E., Toth, R., Nagy, F., et al. (2005). Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* 309, 630–633. doi: 10.1126/science.1115581

Dunlap, J. C. (1999). Molecular bases for circadian clocks. *Cell* 96, 271–290. doi: 10.1016/S0092-8674(00)80566-8

Edery, I. (1999). Role of posttranscriptional regulation in circadian clocks: lessons from *Drosophila*. *Chronobiol. Int.* 16, 377–414. doi: 10.3109/07420529908998716

Felsenstein, J. (1985). Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. *Evolution* 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb00420.x

Goodspeed, D., Chehab, E. W., Min-Venditti, A., Braam, J., and Covington, M. F. (2012). Arabidopsis synchronizes jasmonate-mediated defense with insect circadian behavior. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4674–4677. doi: 10.1073/pnas.1116368109

Goodspeed, D., Liu, J. D., Chehab, E. W., Sheng, Z. J., Francisco, M., Kliebenstein, D. J., et al. (2013). Postharvest circadian entrainment enhances crop pest resistance and phytochemical cycling. *Curr. Biol.* 23, 1235–1241. doi: 10.1016/j.cub.2013.05.034

Graf, A., Schlereth, A., Stitt, M., and Smith, A. M. (2010). Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proc. Natl. Acad. Sci. U.S.A.* 107, 9458–9463. doi: 10.1073/pnas.0914299107

Green, R. M., Tingay, S., Wang, Z. Y., and Tobin, E. M. (2002). Circadian rhythms confer a higher level of fitness to Arabidopsis plants. *Plant Physiol.* 129, 576–584. doi: 10.1104/pp.004374

Gutierrez, R. A., Stokes, T. L., Thum, K., Xu, X., Obertello, M., Katari, M. S., et al. (2008). Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene *CCA1*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4939–4944. doi: 10.1073/pnas.0800211105

Harmer, S. L. (2009). The circadian system in higher plants. *Annu. Rev. Plant Biol.* 60, 357–377. doi: 10.1146/annurev.arplant.043008.092054

Haydon, M. J., Hearn, T. J., Bell, L. J., Hannah, M. A., and Webb, A. A. (2013). Metabolic regulation of circadian clocks. *Semin. Cell Dev. Biol.* 24, 414–421. doi: 10.1016/j.semcdb.2013.03.007

Hess, J. R., Carman, J. G., and Banowetz, G. M. (2002). Hormones in wheat kernels during embryony. *J. Plant Physiol.* 159, 379–386. doi: 10.1016/j.jplph.2004.01.003

Hotta, C. T., Gardner, M. J., Hubbard, K. E., Baek, S. J., Dalchau, N., Suhita, D., et al. (2007). Modulation of environmental responses of plants by circadian clocks. *Plant Cell Environ.* 30, 333–349. doi: 10.1111/j.1365-3040.2006.01627.x

Izawa, T., Mihara, M., Suzuki, Y., Gupta, M., Itoh, H., Nagano, A. J., et al. (2011). Os-GIGANTEA confers robust diurnal rhythms on the global transcriptome of rice in the field. *Plant Cell* 23, 1741–1755. doi: 10.1105/tpc.111.083238

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074–1080. doi: 10.1126/science.1063127

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275

Ko, D. K., Rohozinski, D., Song, Q., Taylor, S. H., Juenger, T. E., Harmon, F. G., et al. (2016). Temporal shift of circadian-mediated gene expression and carbon fixation contributes to biomass Heterosis in Maize Hybrids. *PLoS Genet.* 12:e1006197. doi: 10.1371/journal.pgen.1006197

Kong, Y. M., Han, L., Liu, X., Wang, H. F., Wen, L. Z., Yu, X. L., et al. (2020). The nodulation and nyctinastic leaf movement is orchestrated by clock gene *LHY* in *Medicago truncatula*. *J. Integr. Plant Biol.* 62, 1880–1894. doi: 10.1111/jipb.12999

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

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

Legnaioli, T., Cuevas, J., and Mas, P. (2009). TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. *EMBO J.* 28, 3745–3757. doi: 10.1038/emboj.2009.297

Letunic, I., and Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* 46, D493–D496. doi: 10.1093/nar/gkx922

Li, N., Wei, S., Chen, J., Yang, F., Kong, L., Chen, C., et al. (2018). OsASR2 regulates the expression of a defence-related gene, Os2H16, by targeting the GT-1 cis-element. *Plant Biotechnol. J.* 16, 771–783. doi: 10.1111/pbi.12827

Li, X., and Xu, K. (2014). Effects of exogenous hormones on leaf photosynthesis of *Panax ginseng*. *Photosynthetica* 52, 152–156. doi: 10.1007/s11099-014-0005-1

Li, Y., Fu, X., Zhao, M., Zhang, W., Li, B., An, D., et al. (2018). A genome-wide view of transcriptome dynamics during early spike development in bread wheat. *Sci. Rep.* 8:15338. doi: 10.1038/s41598-018-33718-y

Li, Z., Wang, M., Lin, K., Xie, Y., Guo, J., Ye, L., et al. (2019). The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol.* 20:139. doi: 10.1186/s13059-019-1746-8

Lin, Y. P., Lee, T. Y., Tanaka, A., and Charng, Y. Y. (2014). Analysis of an Arabidopsis heat-sensitive mutant reveals that chlorophyll synthase is involved in reutilization of chlorophyllide during chlorophyll turnover. *Plant J.* 80, 14–26. doi: 10.1111/tpj.12611

Lowrey, P. L., and Takahashi, J. S. (2004). Mammalian circadian biology: elucidating genome-wide levels of temporal organization. *Annu. Rev. Genomics Hum. Genet.* 5, 407–441. doi: 10.1146/annurev.genom.5.061903.175925

Lu, S. X., Knowles, S. M., Andronis, C., Ong, M. S., and Tobin, E. M. (2009). CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL function synergistically in the circadian clock of Arabidopsis. *Plant Physiol* 150, 834–843. doi: 10.1104/pp.108.133272

Malapeira, J., Khaitova, L. C., and Mas, P. (2012). Ordered changes in histone modifications at the core of the Arabidopsis circadian clock. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21540–21545. doi: 10.1073/pnas.1217022110

Maloney, V. J., Park, J. Y., Unda, F., and Mansfield, S. D. (2015). Sucrose phosphate synthase and sucrose phosphate phosphatase interact in planta and promote plant growth and biomass accumulation. *J. Exp. Bot.* 66, 4383–4394. doi: 10.1093/jxb/erv101

Masidur Alam, M., Naeem, M., Idrees, M., Masroor, M., Khan, A., and Moinuddin. (2012). Augmentation of photosynthesis, crop productivity, enzyme activities and alkaloids production in Sadabahar (*Catharanthus roseus* L.) through application of diverse plant growth regulators. *J. Crop Sci. Biotechnol.* 15, 117–129. doi: 10.1007/s12892-011-0005-7

Meyer, P., Saez, L., and Young, M. W. (2006). PER-TIM interactions in living Drosophila cells: an interval timer for the circadian clock. *Science* 311, 226–229. doi: 10.1126/science.1118126

Ni, Z., Kim, E. D., Ha, M., Lackey, E., Liu, J., Zhang, Y., et al. (2009). Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457, 327–331. doi: 10.1038/nature07523

Nozue, K., Covington, M. F., Duek, P. D., Lorrain, S., Fankhauser, C., Harmer, S. L., et al. (2007). Rhythmic growth explained by coincidence between internal and external cues. *Nature* 448, 358–361. doi: 10.1038/nature05946

Perales, M., and Mas, P. (2007). A functional link between rhythmic changes in chromatin structure and the Arabidopsis biological clock. *Plant Cell* 19, 2111–2123. doi: 10.1105/tpc.107.050807

Ramirez-Gonzalez, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361:eaar6089. doi: 10.1126/science.aar6089

Reinbothe, S., Reinbothe, C., Lebedev, N., and Apel, K. (1996). PORA and PORB, two light-dependent protochlorophyllide-reducing enzymes of angiosperm chlorophyll biosynthesis. *Plant Cell* 8, 763–769. doi: 10.1105/tpc.8.5.763

Reppert, S. M., and Weaver, D. R. (2002). Coordination of circadian timing in mammals. *Nature* 418, 935–941. doi: 10.1038/nature00965

Saitou, N., and Nei, M. (1987). The neighbor-joining method - a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sanchez, S. E., and Kay, S. A. (2016). The plant circadian clock: from a simple timekeeper to a complex developmental manager. *Cold Spring Harb. Perspect. Biol.* 8:a027748. doi: 10.1101/cshperspect.a027748

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I. A., et al. (1998). The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93, 1219–1229. doi: 10.1016/s0092-8674(00)81465-8

Shan, Q., Wang, Y., Li, J., and Gao, C. (2014). Genome editing in rice and wheat using the CRISPR/Cas system. *Nat. Protoc.* 9, 2395–2410. doi: 10.1038/nprot.2014.157

Song, H. R., and Noh, Y. S. (2012). Rhythmic oscillation of histone acetylation and methylation at the Arabidopsis central clock loci. *Mol. Cells* 34, 279–287. doi: 10.1007/s10059-012-0103-5

Steed, G., Ramirez, D. C., Hannah, M. A., and Webb, A. A. (2021). Chronoculture, harnessing the circadian clock to improve crop yield and sustainability. *Science* 372:eabc9141. doi: 10.1126/science.abc9141

Stitt, M., and Zeeman, S. C. (2012). Starch turnover: pathways, regulation and role in growth. *Curr. Opin. Plant Biol.* 15, 282–292. doi: 10.1016/j.pbi.2012.03.016

Strayer, C., Oyama, T., Schultz, T. F., Raman, R., Somers, D. E., Mas, P., et al. (2000). Cloning of the Arabidopsis clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* 289, 768–771. doi: 10.1126/science.289.5480.768

Sun, C., Zhang, K., Zhou, Y., Xiang, L., He, C., Zhong, C., et al. (2021). Dual function of clock component OsLHY sets critical day length for photoperiodic flowering in rice. *Plant Biotechnol. J.* 19, 1644–1657. doi: 10.1111/pbi.13580

Wang, F., Han, T., Song, Q., Ye, W., Song, X., Chu, J., et al. (2020). The Rice Circadian Clock Regulates Tiller Growth and Panicle Development Through Strigolactone Signaling and Sugar Sensing. *Plant Cell* 32, 3124–3138. doi: 10.1105/tpc.20.00289

Wang, K., Bu, T., Cheng, Q., Dong, L., Su, T., Chen, Z., et al. (2021). Two homologous LHY pairs negatively control soybean drought tolerance by repressing the abscisic acid responses. *New Phytol.* 229, 2660–2675.

Wang, W., Barnaby, J. Y., Tada, Y., Li, H., Tor, M., Caldelari, D., et al. (2011). Timing of plant immune responses by a central circadian regulator. *Nature* 470, 110–114. doi: 10.1038/nature09766

Wang, Z. Y., and Tobin, E. M. (1998). Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93, 1207–1217. doi: 10.1016/s0092-8674(00)81464-6

Wu, Z., Zhang, X., He, B., Diao, L., Sheng, S., Wang, J., et al. (2007). A chlorophyll-deficient rice mutant with impaired chlorophyllide esterification in chlorophyll biosynthesis. *Plant Physiol.* 145, 29–40. doi: 10.1104/pp.107.100321

Yakir, E., Hilman, D., Kron, I., Hassidim, M., Melamed-Book, N., and Green, R. M. (2009). Posttranslational regulation of CIRCADIAN CLOCK ASSOCIATED1 in the circadian oscillator of Arabidopsis. *Plant Physiol.* 150, 844–857. doi: 10.1104/pp.109.137414

Yamori, W., and Shikanai, T. (2016). Physiological functions of cyclic electron transport around photosystem i in sustaining photosynthesis and plant growth. *Annu. Rev. Plant Biol.* 67, 81–106.

Yang, R., Li, P., Mei, H., Wang, D., Sun, J., Yang, C., et al. (2019). Fine-tuning of MiR528 accumulation modulates flowering time in rice. *Mol. Plant* 12, 1103–1113. doi: 10.1016/j.molp.2019.04.009

Yanovsky, M. J., Izaguirre, M., Wagmaister, J. A., Gatz, C., Jackson, S. D., Thomas, B., et al. (2000). Phytochrome A resets the circadian clock and delays tuber formation under long days in potato. *Plant J.* 23, 223–232. doi: 10.1046/j.1365-313x.2000.00775.x

Young, M. W., and Kay, S. A. (2001). Time zones: a comparative genetics of circadian clocks. *Nat. Rev. Genet.* 2, 702–715.

Yue, W., Fulai, S., Qingrong, G., Yanxia, Z., Nan, W., and Weidong, Z. (2017). Auxins regulations of branched spike development and expression of *TFL*, a LEAFY-Like Gene in Branched Spike Wheat (*Triticum aestivum*). *J. Agric. Sci.* 9, 27–36.

Zhang, Z., Chen, J., Su, Y., Liu, H., Chen, Y., Luo, P., et al. (2015). TaLHY, a 1R-MYB transcription factor, plays an important role in disease resistance against stripe rust fungus and ear heading in wheat. *PLoS One* 10:e0127723. doi: 10.1371/journal.pone.0127723

Zhou, X., and Su, Z. (2007). EasyGO: Gene Ontology-based annotation and functional enrichment analysis tool for agronomical species. *BMC Genomics* 8:246. doi: 10.1186/1471-2164-8-246

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 25 July 2022
DOI 10.3389/fpls.2022.929747

OPEN ACCESS

EDITED BY
Fanjiang Kong,
Guangzhou University,
China

REVIEWED BY
José Antonio Abelenda,
National Institute of Agricultural and Food Research and Technology, Spain
Lin Zhao,
Northeast Agricultural University,
China

*CORRESPONDENCE
Tianfu Han
hantianfu@caas.cn
Lijie Yu
yulijie1961@126.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 27 April 2022
ACCEPTED 04 July 2022
PUBLISHED 25 July 2022

CITATION
Yuan S, Wang Y, Wang J, Zhang C, Zhang L, Jiang B, Wu T, Chen L, Xu X, Cai Y, Sun S, Chen F, Song W, Wu C, Hou W, Yu L and Han T (2022) *GmFT3a* fine-tunes flowering time and improves adaptation of soybean to higher latitudes.
Front. Plant Sci. 13:929747.
doi: 10.3389/fpls.2022.929747

GmFT3a fine-tunes flowering time and improves adaptation of soybean to higher latitudes

Shan Yuan[1†], Yining Wang[1,2†], Junya Wang[1,2], Chunlei Zhang[1,2], Lixin Zhang[1], Bingjun Jiang[1], Tingting Wu[1], Li Chen[1], Xin Xu[1], Yupeng Cai[1], Shi Sun[1], Fulu Chen[1], Wenwen Song[1], Cunxiang Wu[1], Wensheng Hou[1], Lijie Yu[2]* and Tianfu Han[1,2]*

[1]MARA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [2]College of Life Science and Technology, Harbin Normal University, Harbin, China

Onset of flowering of plants is precisely controlled by extensive environmental factors and internal molecular networks, in which *FLOWERING LOCUS T* (*FT*) is a key flowering integrator. In soybean, a typical short-day plant, 11 *FT* homologues are found in its genome, of which several homologues are functionally diversified in flowering pathways and the others including *GmFT3a* are yet unknown. In the current study, we characterized *GmFT3a*, which is located on the same chromosome as the flowering promoters *GmFT2a* and *GmFT5a*. Overexpression of *GmFT3a* significantly promoted flowering of Arabidopsis under the inductive long-day (LD) photoperiod. *GmFT3a* over-expressed soybean also flowered earlier than the control under LD, but they were not significantly different under inductive short-day (SD) conditions, indicating that *GmFT3a* acts as a flowering promoter in the non-inductive photoperiod in soybean. Compared with other *GmFT* homologues, *GmFT3a* exhibited a slighter effect in flowering promotion than *GmFT2a*, *GmFT5a* and *GmFT2b* under LD conditions. *GmFT3a* promoted flowering by regulating the expression of downstream flowering-related genes and also affected the expression of other *GmFTs*. According to the re-sequencing data, the regional distributions of two major haplotypes in 176 soybean varieties were analyzed. The varieties with *GmFT3a*-Hap2 haplotype matured relatively early, and relative higher expression of *GmFT3a* was detected in early maturing varieties, implying that Hap2 variation may contribute to the adaptation of soybean to higher latitude regions by increasing expression level of genes in metabolism and signaling pathways. The early flowering germplasm generated by overexpression of *GmFT3a* has potential to be planted at higher latitudes where non-inductive long day is dominant in the growing season, and *GmFT3a* can be used to fine-tune soybean flowering and maturity time and improve the geographical adaptation.

KEYWORDS

soybean, *GmFT3a*, flowering time, photoperiod, adaptation

Introduction

Precise timing of flowering is critical to the environmental adaptation and productivity of crops (Cockram et al., 2007; Nakamichi, 2015). Understanding the molecular mechanisms underlying reproductive transition is a prerequisite for improving regional adaptability in crop breeding (Blümel et al., 2015; Lin et al., 2021a). In Arabidopsis, photoperiod, vernalization, gibberellin (GA), age and autonomous pathways were confirmed to be involved in the modulation of flowering (Zeevaart, 2008; Blümel et al., 2015). Among the environmental factors, photoperiod acts as a major determinant signal for flowering in many plants (Thomas and Vince-Prue, 1997; He et al., 2020). In the photoperiodic pathway, *FLOWERING LOCUS T* (*FT*), a key photoperiod-regulated flowering integrator, encodes florigen which functions as a leaf-derived long-distance mobile signals and promotes floral transition (Corbesier et al., 2007; Wigge, 2011). Homologues of *FT* are highly conserved and promotive to flowering in diverse species (Tamaki et al., 2007; Pin et al., 2010; Park et al., 2014; Qin et al., 2017; Zhang et al., 2021).

Soybean is a typical short-day plant but is now widely grown in a wide range of latitudes from 53°N to 40°S with diverse daylength (Watanabe et al., 2012; Jia et al., 2014; Liu et al., 2022), resulting from rich varietal diversity in maturity. However, the cultivation area of a given variety is restricted to a narrow range of latitudes because of its sensitivity to photoperiods. In soybean breeding programs, the maturity or growth period which is mainly controlled by the photoperiod and temperature, is one of the most important traits for adaptation to a given environment (Hartwig, 1970; Wu et al., 2015; Song et al., 2019; Lin et al., 2021a).

Soybean maturity ecotypes were classified into a numerical and consecutive system "Maturity groups (MGs)" according to photothermal sensitivity and adaptation to specific environments (Hartwig, 1970; Jia et al., 2014; Mourtzinis and Conley, 2017; Song et al., 2019). The diversity of soybean varieties in MG and adaptation to different environments benefit from the variations and combinations of the genes in the photoperiod pathway (Jiang et al., 2014; Wang et al., 2015; Zhang et al., 2020; Liu et al., 2020a; Lin et al., 2021b). It is important to deeply understand the effect of allelic variations in maturity-related genes on the adaptation of soybean varieties to diverse geographic regions and farming systems. Soybean is a diploid species derived from an ancient tetraploid, and its genome has undergone whole-genome duplications during its long evolutionary history (Wang et al., 2015). Since then, many genes have multiple copies in the genome. For *FT* homologues, the integrators in the flowering pathway, there are at least 11 homologues in soybean (Kong et al., 2010; Liu et al., 2020b; Lee et al., 2021; Zhang et al., 2021). Among these genes, *GmFT2a* (*Glyma.16g150700*) and *GmFT5a* (*Glyma.16g044100*) play redundant but coordinated roles in flowering promotion by up-regulating the expression of floral determination genes (Kong et al., 2010; Sun et al., 2011; Na et al., 2013; Nan et al., 2014; Takeshima et al., 2019; Lin et al., 2021b). Moreover, *GmFT2b*, promotes flowering under long-days (LDs; Chen et al., 2020). *GmFT2b* overexpression plants exhibited early flowering under non-inductive LD conditions, while *ft2b* mutants flowered later than wildtype (WT) under LD. However, *GmFT1a* and *GmFT4* were identified as inhibiting genes for flowering. Overexpression of *GmFT1a* in soybean delayed flowering, confirming that it was a flowering repressor in soybean (Liu et al., 2018), and *GmFT4* was proven to be a repressor in Arabidopsis (Zhai et al., 2014). *GmFT3b* acts redundantly in flowering time regulation and may be compensated by other *FT* homologs in soybean (Su et al., 2022). The existence of multiple soybean *GmFT* homologues might enhance the adjustability of photoperiodic regulation of flowering (Jiang et al., 2019; Lin et al., 2021b). For instance, the expression of FT-like proteins is not restricted to short-day (SD) conditions (Kobayashi and Weigel, 2007), highlighting the coincident expression of FT-like genes encoding both floral activators and floral inhibitors in the day-neutral species of *Solanum lycopersicum* (Cao et al., 2016a) and *Nicotiana tabacum* (Beinecke et al., 2018). However, some important *GmFT* homologues have yet to be elucidated, which hinders a better understanding of the mechanism of functional allocations of the *GmFT* family in the photoperiodic flowering pathway. In this study, we focused on the *GmFT3a* (*Glyma.16g044200*) which is highly homologous to *GmFT3b*, and investigated its function in the flowering control of soybean. These results will deepen the understanding of the functional diversification of *GmFT* members and provide a new target gene for improving the adaptation of soybean to diverse regions.

Materials and methods

Plant materials and growth conditions

For gene cloning of *GmFT3a*, Zigongdongdou (ZGDD), a late-maturing (MG IX) and photoperiod-sensitive soybean [*Glycine max* (L.) Merr.] variety which originated from Sichuan province in the southwest China (Wu et al., 2006) was selected as the plant material. Heihe 27 (HH27), an early maturing (MG 00) and photoperiodic-insensitive and early maturing variety (MG 00) from Heilongjiang province in northeast China, was selected to compare the gene expression pattern with that of ZGDD. For genetic transformation, a medium-maturing (MG III) variety Jack was used as the receptor. Moreover, 12 MG standard soybean varieties were chosen for gene expression analysis of *GmFT3a*. All of soybean seeds were germinated in pots with soil and vermiculite (1:1) and the plants were grown in growth chambers with constant temperature (28°C) and controlled photoperiods, i.e., LD (16 h light/8 h dark) and SD (12 h light/12 h dark), respectively. A diverse panel of 176 soybean cultivars with consecutive MG groups covering early MG 000 to MG IX, were re-sequenced to investigate the natural variation of *GmFT3a*. The germplasms were collected from China and the United States (Supplementary Table S2; Zhang et al., 2019; Liu et al., 2020c).

GmFT3a gene cloning

Total RNA was extracted for gene cloning and expression analysis from the first trifoliolate leaf of ZGDD in LD (16 h light/8 h dark) using TransZol UP kit (TransGen Biotech, Beijing, China). The first strand of cDNA was obtained using the TransGene reverse transcription kit (TransGen Biotech, Beijing, China). The corresponding primer sequences are listed in Supplementary Table S1.

Gene expression analysis

ZGDD, HH27 and 12 maturity group (MG) standard varieties (Supplementary Table S2) were treated with LD and SD, respectively. After photoperiodic treatments (LD and SD) for 13 days, various tissues (shoot apex, unifoliolate, trifoliolate, stem, hypocotyl, root and flower) were taken to evaluate the expression level of *GmFT3a*. The real-time quantitative PCR (qRT-PCR) primers were designed by NCBI Primer Blast online with the amplification products ranging from 100 to 300 bp. Three-step PCR method was set with *GmActin* as the internal reference on the ABI Prism® 7900HT real-time PCR instrument, and that relative expression of target genes was calculated according to the $2^{-\Delta\Delta CT}$ algorithm (Pin and Nilsson, 2012).

Subcellular localization of GmFT3a protein

The subcellular localization of GmFT3a protein was conducted by constructing p16318-*GmFT3a*-GFP fusion expression vector and carrying out plasmolysis isolation, transformation and culture of Arabidopsis protoplasts. The fusion *GmFT3a*-p16318- GFP subcellular localization vector of the experimental group and the empty p16318 vector of the control group were prepared and transformed into Arabidopsis protoplasts, and the expression position of GFP was observed in confocal laser scanning microscope (Olympus FV31-SD, Olympus Corporation, Japan).

Expression vector construction and plant transformation

A fragment with *Xba*I and *Kpn*I digestion sites was ligated into pCAMBIA1300 and pTF101.1 vector to construct overexpression vectors, and the constructed vectors were transformed into Arabidopsis Columbia-0 by *Agrobacterium tumefaciens* mediation and inflorescence dipping method (Clough and Bent, 1998). The soybean variety Jack was used for transformation according to a previously published protocol (Chen et al., 2018). The transformed Arabidopsis plants were screened by hygromycin (Hyg), and the soybean plants were screened by applying 160 mg/l glufosinate solution on young leaves, and PCR identification.

Transcriptome sequencing and data analysis

Leaf was sampled at 13 days after emergence (DAE) under SD conditions and 30 DAE under LD conditions. Three biological replicates were analyzed. The total amount of 1.5 μg RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, United States) following manufacturer's manual. In order to select cDNA fragments of 150 ~ 200 bp in length preferentially, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, United States). The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. The raw data have been uploaded to NCBI Sequence Read Archive with the accession number of PRJNA832118.

All the downstream analyses were based on clean data with high quality. Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences); Pfam (Protein family); KOG/COG (Clusters of Orthologous Groups of proteins); Swiss-Prot (A manually annotated and reviewed protein sequence database); KO (KEGG Ortholog database); GO (Gene Ontology). Differential expression analysis of two conditions/groups was performed using the DESeq R package (1.10.1; Li and Dewey, 2011). DESeq provided statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. Genes with an adjusted p value <0.05 found by DESeq were assigned as differentially expressed. Gene Ontology (GO) enrichment analysis of the differentially expressed genes (DEGs) was implemented by the GOseq R packages-based Wallenius non-central hyper-geometric distribution (Young et al., 2010), which can adjust for gene length bias in DEGs. We used KOBAS (Mao et al., 2005) software to test the statistical enrichment of differential expression genes in KEGG pathways.

Phenotyping and statistical analysis

Transgenic plants were grown in culture rooms under SD and LD conditions, respectively, and the flowering time was determined as the number of days from VE (emergence) to R1 stage (beginning bloom: the first flower appears at any node in the main stem; Fehr and Caviness, 1977). The statistical analysis was carried out by Microsoft Excel, and Student's t-test was used to assess the significance of difference between the transgenic lines.

Natural variations and haplotype identification

In order to evaluate the effects of *GmFT3a* natural variations on the adaptability of soybean varieties, the *GmFT3a* alleles of 176 soybeans germplasm were detected. The genomic data were obtained in our previous investigation (Zhang et al., 2019; Liu et al., 2020c) and the MG classification were carried out according to the standard methodology described by Song et al. (2019).

Results

Bioinformatic characteristics and subcellular localization of GmFT3a

The length of the *GmFT3a* CDS region is 528 bp in ZGDD, encoding 175 amino acids, which is a hydrophilic non-secretory protein. The region ranging from amino acid 27 to 164 is the PEBP domain (Supplementary Figures S1, S2). The phylogenetic tree divided these proteins into three clear groups based on their protein sequence homology, including (1) FT, (2) MFT (MOTHER of FT and TFL 1) and (3) BFT (BROTHER of FT and TFL 1) and TFL 1 (TERMINAL FLOWER 1) clades. GmFT3a was closely clustered with AtFT (AT1G65480.1; Figure 1A). The tertiary structure of its protein is similar to that of the AtFT protein and GmFT2a protein (Supplementary Figure S3). p16318-GFP in the control group did not show subcellular expression specificity, while fluorescence signals of GmFT3a-GFP were detected in cytoplasm and nucleus (Figure 1B). This result is similar to the subcellular localization of other functional phosphatidylethanolamine-binding proteins in soybean (Wang et al., 2015).

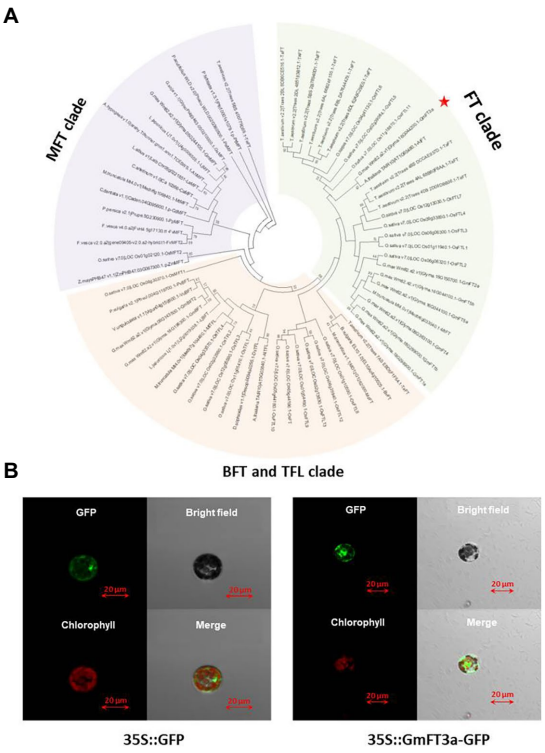


FIGURE 1
Phylogenetic tree based on protein sequences between *GmFT3a* and other FT/TFL1 family members from flowering plants **(A)** and subcellular localization of GmFT3a-GFP fusion protein in Arabidopsis protoplasts. The left four panels are 35S::GmFT3a-GFP constructs, and the right four panels are controls (35S::GFP) **(B)**.

GmFT3a expression was regulated by photoperiods and correlated with maturity groups of varieties

GmFT3a expression was evaluated in different tissues of the photoperiod-sensitive variety ZGDD. Under both SD and LD conditions, relatively higher expression was detected in trifoliolate, followed by the shoot apex and unifoliolate, and extremely lower expression was detected in stem, hypocotyl, root and flower tissues in ZGDD (Supplementary Figure S4a). The expression of *GmFT3a* was not significantly different among trifoliolate, shoot apex, and unifoliolate under SD, while the expression of trifoliolate was significantly higher than that of shoot apex and unifoliolate under LD and exceeded the average expression under SD (Supplementary Figure S4a).

We also analyzed the diurnal expression patterns of *GmFT3a* in trifoliolates of both HH27 and ZGDD since 13 DAE. In early variety HH27, and the consecutive 48 h sampling by every 4 h showed two major expression peaks at 4 h and 32 h in SD condition, but relative lower expression under LD, and only one peak on 0/48 h (Supplementary Figure S4b). In late variety ZGDD, the consecutive 48 h sampling every 4 h showed two peaks at 40 h and 4 h under LD, and no significant diurnal rhythm was displayed under SD (Supplementary Figure S4b). Under LD condition, the *GmFT3a* expression pattern of ZGDD was similar as under SD, but the highest expression was detected in the unifoliolate value in HH27, then followed by the trifoliolate and shoot apex (Supplementary Figure S4a).

Among the varieties belonging to different maturity groups, it was found that the expression of *GmFT3a* in the trifoliolates of early maturing groups of MG 000, MG 0 and MG VII was higher than that of the late-maturity groups under LD, while MG 000 and MG 00 maintained higher *GmFT3a* expression than the others under SD conditions (Supplementary Figure S4c).

Overexpression of *GmFT3a* promoted flowering and regulated down-stream genes in both Arabidopsis and soybean

A total of four independent transgenic lines overexpressing *GmFT3a* were obtained in Arabidopsis, and an the obvious early flowering phenotype was observed in the T_5 homozygous

generation under LD treatment (Figure 2A). The average flowering time of the four lines was 3.9 d earlier than that of WT Arabidopsis ($p < 0.01$, Figure 2B). The average total (rosette and cauline) leaves of WT Arabidopsis were 16.7 ± 0.6, and those of the *GmFT3a* Arabidopsis overexpressing lines were 11.2 ± 1.0, 12.6 ± 1.3, 10.7 ± 1.4 and 12.5 ± 0.7, respectively. The average total leaf number of four lines was reduced by 29.6% compared with that of the WT ($p < 0.01$). Overexpression of *GmFT3a* resulted in a very significant early flowering phenotype ($p < 0.01$;s Figure 2C) in Arabidopsis. The qRT-PCR results showed that the ectopic *GmFT3a* caused significantly high expression of *AtFT*, *AtSOC1* and *AtCO* ($p < 0.01$, Figure 2D), and the up-regulation of the *AtFT* and *AtCO* in the photoperiod pathway might be the underlying reason for early flowering of transgenic *GmFT3a* Arabidopsis.

The fusion expression vector of *GmFT3a* was constructed and the variety Jack was used for *Agrobacterium*-mediated transformation. According to the flowering time of T_1 generation and the number of seeds harvested, we selected three (Line 7, Line 9 and Line 14) from the *GmFT3a* overexpression lines (Figure 3A). In the T_2 generation, the average flowering time of the Jack variety was 23.7 ± 1.5 DAE under SD, and that of the three transgenic *GmFT3a* lines were 21.7 ± 3.3, 21.5 ± 1.0 and 22.6 ± 3.7 DAE, respectively (Figure 3B). Under LD, the average flowering time of WT was 51.7 ± 4.1 DAE, while the three transgenic *GmFT3a* lines were 8.9 d, 4.0 d and 5.6 d earlier than that the WT, respectively (Figures 3B,C).

The expression of *GmFT* homologues was measured in the transgenic *GmFT3a* soybean under LD condition, and their influence on other *GmFT* genes was analyzed. The results showed that the expression level of transgenic *GmFT3a* plants was significantly higher than that of WT ($p < 0.01$), and the expression level of *GmFT3b*, which has the highest homology with *GmFT3a*, was also significantly higher ($p < 0.01$), while the expression of *GmFT2a* and *GmFT5a* decreased significantly in the transgenic *GmFT3a* soybean ($p < 0.01$). The declining degree of *GmFT2a* expression was higher than that of *GmFT5a*. The expression of *GmFT1a* significantly decreased, while *GmFT1b* increased significantly in transgenic Line 9 and Line 14 ($p < 0.01$; Figure 4).

More genes in metabolism and signaling pathways were enriched under LD than under SD among DEGs in transgenic *GmFT3a* soybean

According to the transcriptomic profiles, a total of 3,129 DEGs with 1906 up-regulated and 1,223 down-regulated unigenes were observed under LD, and 1,001 DEGs including 489 up- and 512 down-regulated unigenes were expressed under SD (Table 1; Figure 5). Among these DEGs, there were 226 (153 up- and 73 down-regulated) unigenes under both LD and SD conditions (Table 1; Figure 5). Among the up-regulated unigenes under LD, 558 up-regulated genes were identified in the transgenic *GmFT3a* soybean ($p < 0.05$). These genes included F-box protein (a gene for the controlled degradation of cellular protein); disease resistance

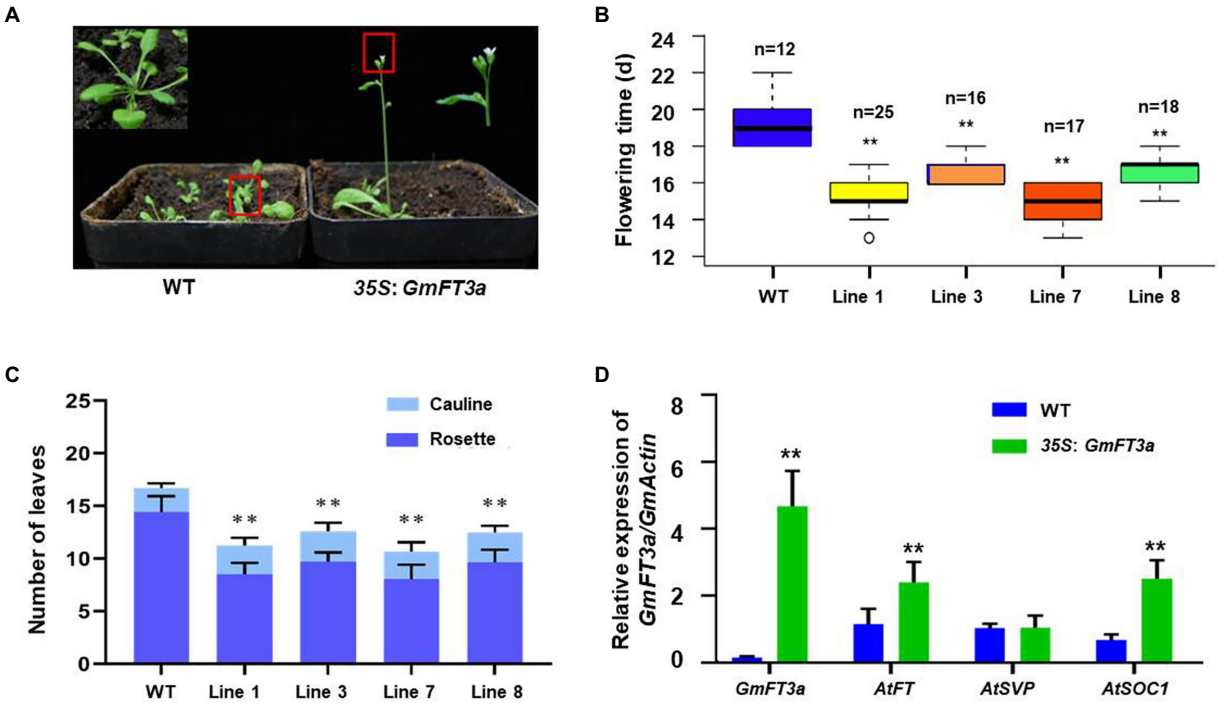


FIGURE 2
The phenotype **(A)**, flowering time **(B)**, the number of rosette and cauline rosette leaves at flowering **(C)** and the expression of flowering regulatory genes **(D)** of the transgenic *GmFT3a* lines under LD condition in Arabidopsis.

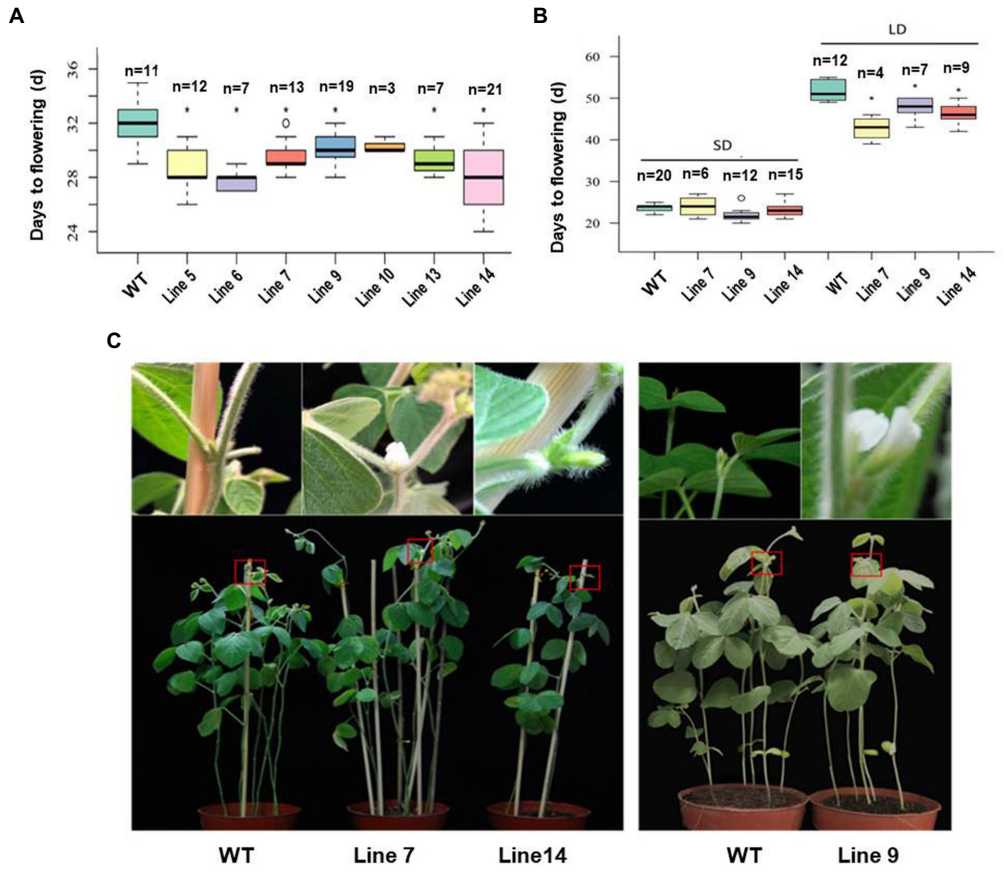


FIGURE 3
The flowering time of the transgenic *GmFT3a* soybean in natural **(A)** and long-day (LD) and short-day (SD) **(B)** conditions and the flowering phenotypes of transgenic *GmFT3a* T_2 generation in LD condition **(C)**. The seeds sown in Beijing, China (N39°58′, E116°20′) on June 20, 2018. The data represent the mean±standard deviation, and statistical significance was determined using Student's *t*-tests (*p<0.05, **p<0.01). DAE: days after emergence.

protein (a resistance protein guard the plant against pathogens); abscisic stress-ripening protein 3 (an abscisic acid-, stress-, and ripening-induced protein); heat shock protein 83 (a gene for promoting maturation), structural maintenance and proper regulation of specific target proteins involved for instance in cell cycle control and signal transduction.

For further identifying the gene function, 49 and 143 GO terms were annotated in the up- and down-regulated DEGs, respectively under LD (Table 2). Under SD condition, only 10 GO terms significantly enriched in the down-regulated DEGs (corrected $p < 0.05$). Among the 49 up-regulated GO items, the "glucosamine-containing compound metabolic process," "fatty acid biosynthetic process," "signal transduction," and "cellular protein modification process" items were significantly overrepresented in biological process; Other terms, such as "proteinaceous extracellular matrix," "extracellular matrix," "extracellular region part" were in cellular component, and "oxidoreductase activity, acting on NAD(P)H, oxygen as acceptor," "kinase activity" and "ATP binding" were in molecular function. While under SD condition, no GO items were significantly enriched in the up-regulated DEGs and only 10 GO items were enriched in the down-regulated DEGs. The top 20 obviously enriched KEGG pathways are shown in Figure 6. By comparing transgenic *GmFT3a* group with control, "Plant hormone signal transduction" and "Amino sugar and nucleotide sugar metabolism" pathway enriched the most in the DEGs under LD condition; while only the "Biosynthesis of secondary metabolites" and "Sesquiterpenoid and triterpenoid biosynthesis" pathway were enriched under the SD condition (Figure 6).

Natural variations of *GmFT3a* existed in soybean varieties with diverse maturity groups and origins

Based on the re-sequencing data, the natural variations of *GmFT3a* in 176 soybean varieties with diverse maturity groups and origins were analyzed and a total of 8 SNPs including 3 alleles in the promoter region (Gm16:4160524.0.4162523) and 5 alleles in genomic region (Gm16:4162524.0.4164824) were identified (Figure 7). Among the two major haplotypes, Hap1 was detected in 169 varieties ranging from MG 000 to MG IX, and Hap2 was

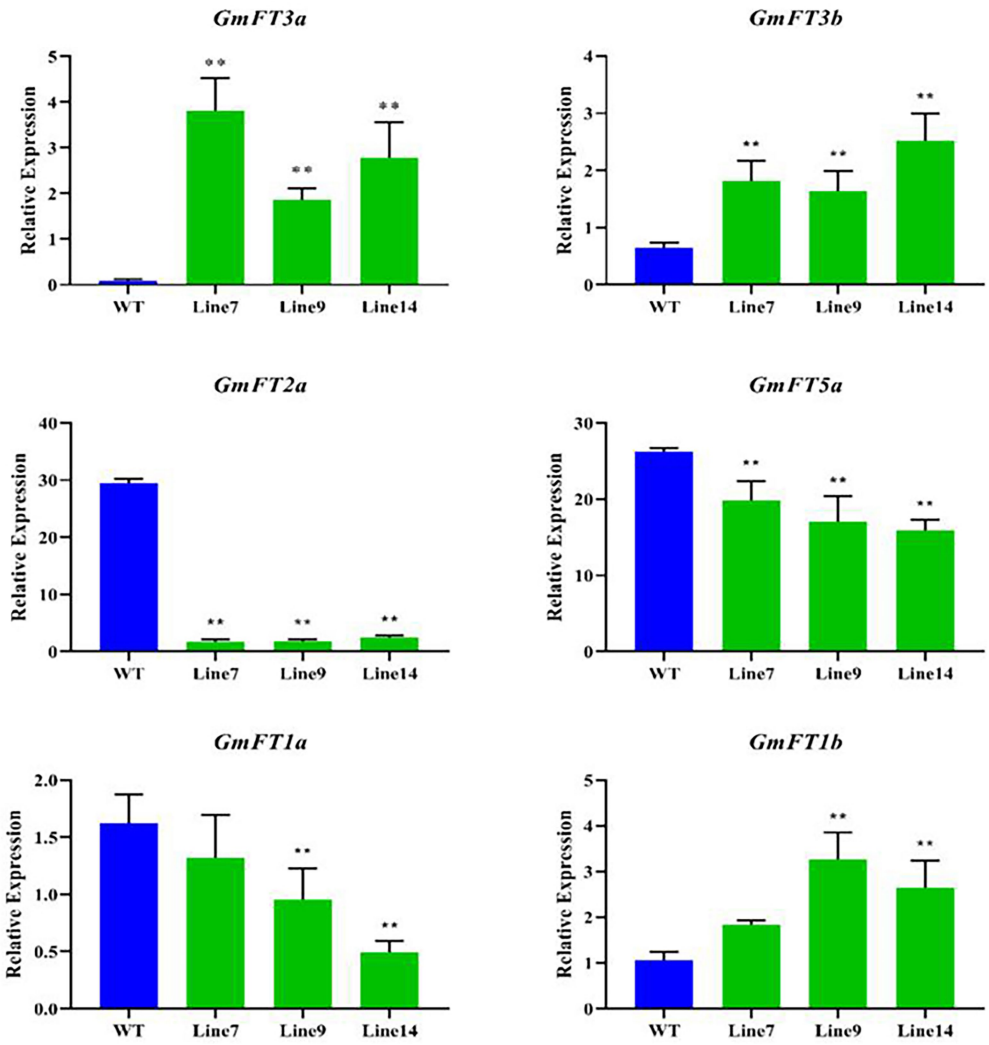


FIGURE 4
Expression analysis of *GmFT3a* in transgenic plants in LD condition.

detected in 7 varieties with MG 00 to MG II (Figure 7; Supplementary Table S2), indicating that Hap2 would be related to early maturity. Almost all Hap2 varieties collected from the high-latitude regions in northeast China, e.g., Heihe 9, Heihe 18, Heihe 38 and Huajiang 4 were bred in Heihe city (approximately 50°N), Heilongjiang Province, and the Tiefeng 18 and Tiefeng 20 varieties were bred in Tieling (approximately 42°N), Liaoning Province in northeast China.

Discussion

GmFT3a fine-tunes flowering time of soybean under LD condition

FT belongs to a phosphatidylethanolamine-binding protein family composed of six members in Arabidopsis (Kardailsky et al., 1999; Jin et al., 2021). FT is considered to be a major component of florigen and plays a critical role in the photoperiodic flowering pathway (Corbesier et al., 2007; Wigge, 2011). To date, the photoperiod-dependent flowering pathway was intensively studied and a total of 11 florigen (*FT*) homologues have been characterized in soybean (Kong et al., 2010; Sun et al., 2011; Liu et al., 2018, 2020b; Zhang et al., 2021; Su et al., 2022). *GmFT3a*, which is located on the same chromosome as the key flowering promoters *GmFT2a* and *GmFT5a*, has yet to be studied for functional identification. In this study, *GmFT3a* promoted flowering by 3.9 d in Arabidopsis under an inductive LD photoperiod, and the overexpression of *GmFT3a* in soybean resulted in earlier flowering significantly by 4.0–8.9 d than that of the control under non-inductive LD conditions (Figures 2A, 3C). Compared with other *GmFT* homologues, *GmFT2a* exhibited a stronger effect in promoting flowering than *GmFT3a*. Overexpression of *GmFT2a* resulted in earlier flowering by

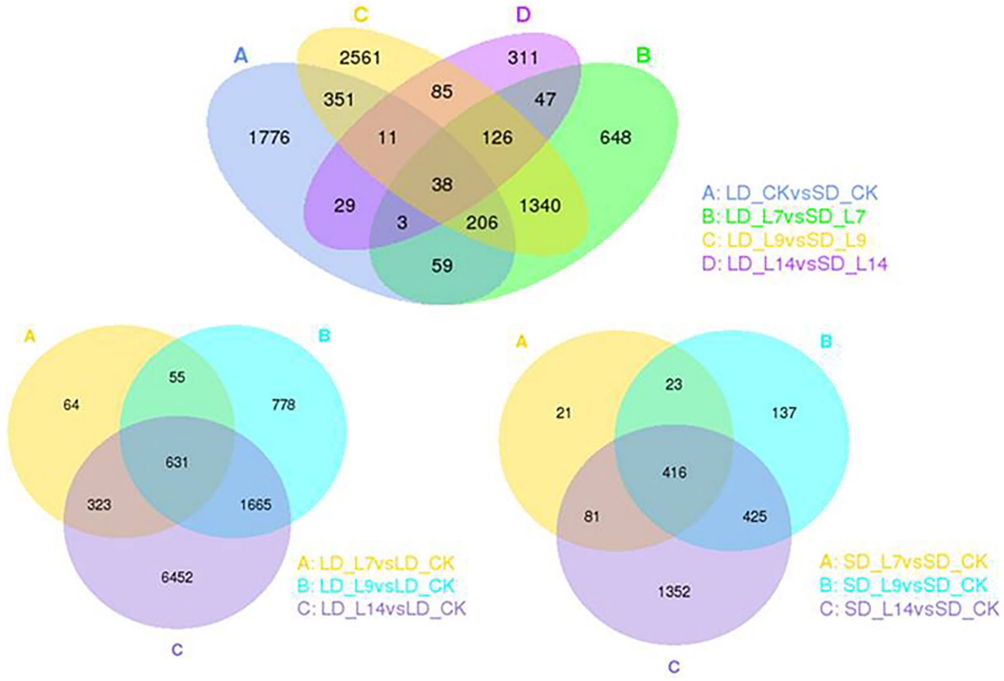


FIGURE 5
The Venn diagram of DEGs between transgenic *GmFT3a* soybean and control in LD and SD, respectively.

approximately 9 d under SD and extremely early flowering phenotypes by approximately 32 d under LD in Jack background (Cai et al., 2020); and under LD, it even drove flowering of the extreme late-maturing variety ZGDD which would have retained its vegetative growth under non-inductive LD conditions (Li et al., 2005; Sun et al., 2011). In addition, *GmFT3a* seems to have a partially similar but weaker function to *GmFT5a* and *GmFT2b*, which held dominant functions under LD condition. The overexpression of *GmFT5a* and *GmFT2b* caused earlier flowering by approximately 16 d and 7 d, respectively under LD (Chen et al., 2018; Cai et al., 2020). Besides, *GmFT3b*, which has the highest homology with *GmFT3a*, was functionally redundant in regulation of flowering time (Jiang et al., 2019; Su et al., 2022). Collectively, the supporting role of *GmFT3a* fine-tuned soybean flowering under the non-inductive LD condition.

GmFT3a promoted flowering by regulating the downstream flowering-related genes

The transition from vegetative growth to reproductive growth is due to the balance of flowering activators and flowering inhibitors according to previous models for photoperiodic flowering in soybean (Liu et al., 2018; Chen et al., 2020). The overexpression of *GmFT3a* unexpectedly down-regulated that of *GmFT2a* and *GmFT5a* under LD (Figure 4), which was inconsistent with the up-regulation of *GmFT2a* and *GmFT5a* by overexpression of *GmFT2b* (Chen et al., 2020). Under LD

TABLE 1 DEGs between transgenic *GmFT3a* soybean and control in LD and SD conditions.

Group	Up	Down	Total
LD	1906	1,223	3,129
SD	489	512	1,001
Specifically differed in LD	1753	1,150	2,903
Specifically differed in SD	354	421	775
Common	153	73	226

conditions, higher expression levels of flowering activators are required to overcome the enhanced the effects of flowering inhibitors. This implied that the function of *GmFT3a* might differ from that of *GmFT2b* in flowering regulation, and the lowered expression of *GmFT2a* and *GmFT5a* is probably responsible for the minor/weak effect on promoting the flowering of *GmFT3a*. *GmFT3b* was up-regulated under the overexpression of *GmFT3a* (Figure 4), and a previous study showed that *GmFT3a* and *GmFT3b* were both significantly up-regulated in the *ft2a* and *ft5a* single soybean mutant, respectively (Li et al., 2021), suggesting that this pair of homologues might share a similar trend under the lack of major promoting effector *GmFT2a* and *GmFT5a*. However, there was no significant change in the expression of *GmFT3a* displayed in both overexpression of *GmFT3b* and the *ft3b* mutant (Su et al., 2022). It could be inferred that *GmFT3a* probably functioned upstream on the *GmFT3b*, and the hypothesis still needs more molecular evidence and flowering phenotype of the *GmFT3a* mutant. In addition, higher expression of *GmFT3a* was detected in the relatively early

TABLE 2 GO items between transgenic *GmFT3a* soybean and control in LD and SD conditions.

Group	Type	Biological process	Cellular component	Molecular function	Total
LD	Up	20	3	26	49
	Down	75	41	27	143
	Total	18	12	29	59
SD	Up	0	0	0	0
	Down	5	2	3	10
	Total	0	1	1	2

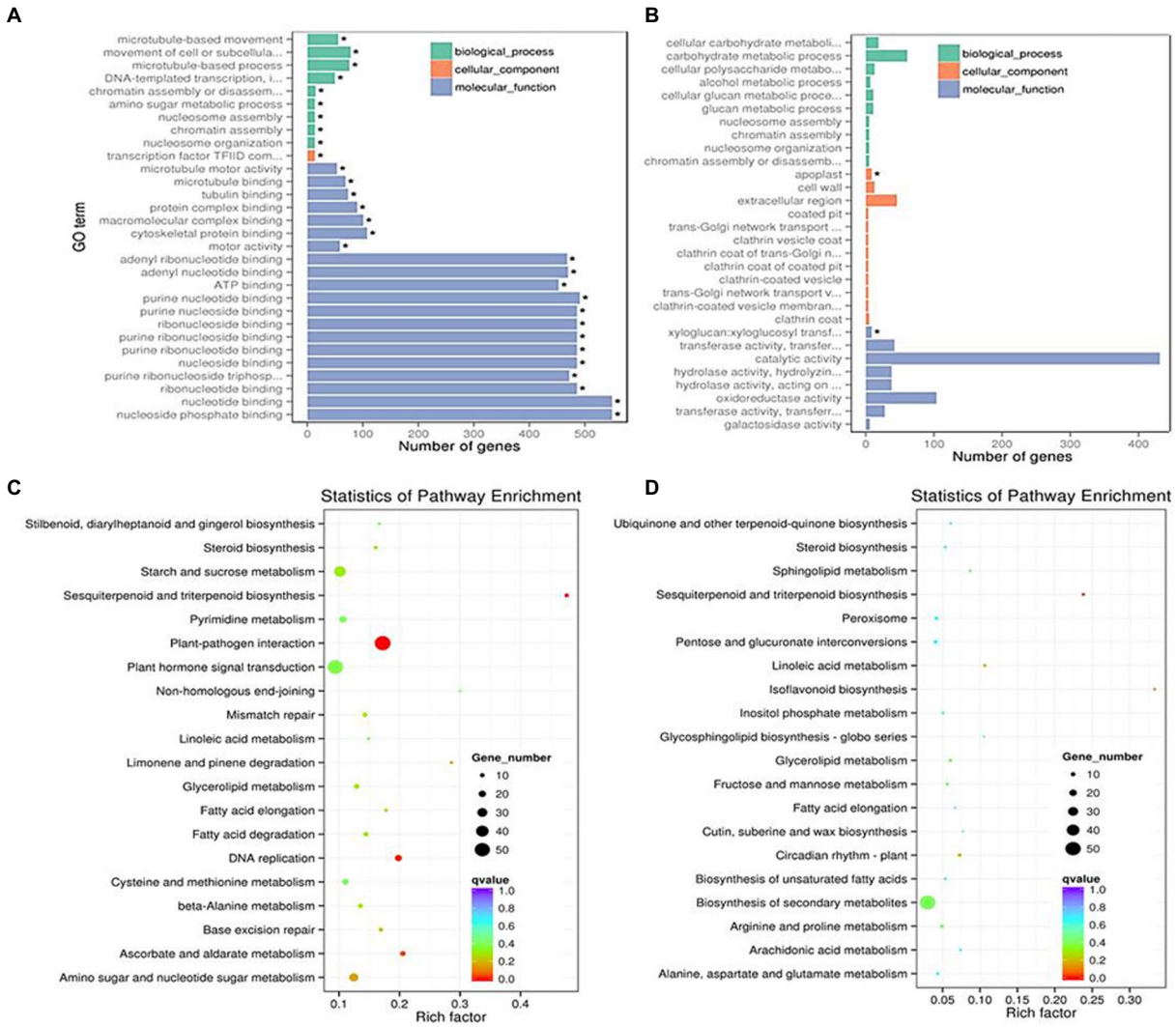


FIGURE 6
The most enriched GO terms in LD **(A)** and SD **(B)** and KEGG pathway terms in LD **(C)** and SD **(D)** conditions.

maturing varieties (MG 000 and MG 00) under SD, suggesting that *GmFT3a* might correlate with early flowering and function as an effective promoter of flowering in non-inductive photoperiodic conditions (Jia et al., 2014; Cao et al., 2016b; Liu et al., 2020b).

In Arabidopsis, the expression of the *AtFT*, *AtSOC1* and *AtCO* genes was significantly increased in the overexpression of *GmFT3a* ($p < 0.05$, Figure 2D). A similar pattern was detected in the overexpression of *GmFT2a* Arabidopsis, which directed the up-regulation of *AtFT* and *AtSOC1* (Sun et al., 2011). The increase in *AtSOC1* in the age pathway was potentially involved in flowering signal integration (Srikanth and Schmid, 2011; Lv et al., 2021). The RNA transcripts of floral identity genes, including *GmAP1*, *GmSOC1*, *GmLFY* and *GmFDL19*, were increased in

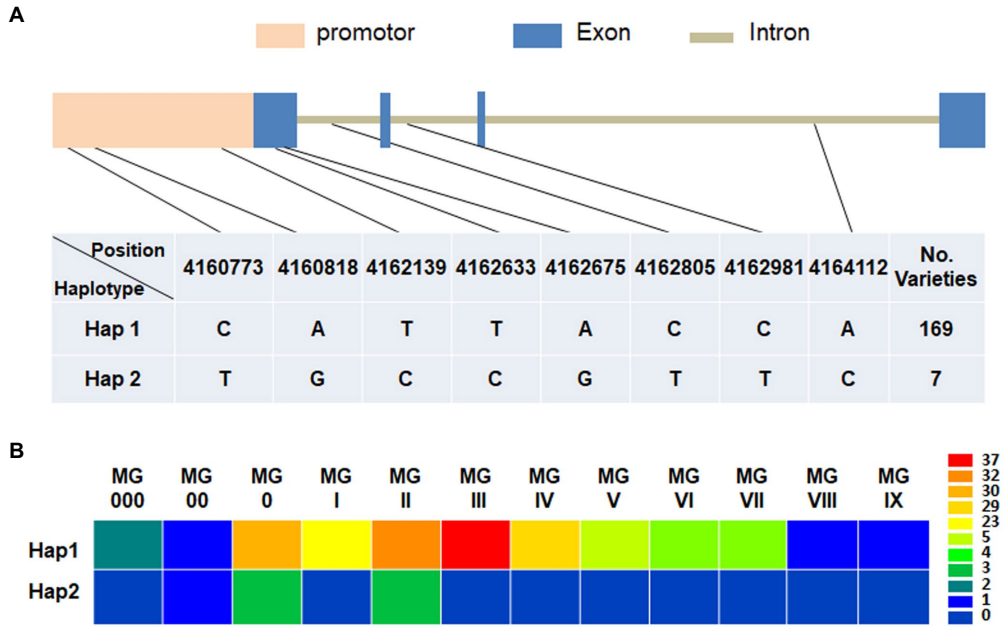


FIGURE 7
Haplotype identification **(A)** and corresponding maturity groups **(B)** of *GmFT3a* in 176 representative soybean varieties.

GmFT2a- or *GmFT5a*-overexpressing soybean lines compared with the wild type (Nan et al., 2014; Takeshima et al., 2019).

Among the transcriptomic DEGs, the large number under LD conditions was more than 3-fold compared to SD conditions (Table 1; Figure 5). The functional patterns of *GmFT3a* were divided into two main aspects: the vegetative growth accelerator and signal transduction pathway. *GmFT3a* probably achieved its promotional roles through the enhancement of genes involved in cell division, photosynthesis, carbohydrate metabolism, fatty acid biosynthesis, and ascorbate metabolism. Moreover, the genes encoding signaling regulation and transcription factors, such as ethylene-responsive transcription factor *AP2-EREBP*, auxin response factor *ARF*, zinc finger protein *COL16*, MADS-box family, and *GRF4* (*GROWTH-REGULATING FACTOR 4*), were involved in plant growth and metabolisms (Supplementary Table S3). These differentially expressed transcription factors were significantly correlated with plant signaling and the regulation of flowering development (Ratcliffe and Riechmann, 2002; Matias-Hernandez et al., 2016; Xu et al., 2021a). Furthermore, the KEGG pathway maps provided abundant information on the exploration of signaling and metabolic pathways involved in *GmFT3a*. Many differentially expressed unigenes were enriched in signaling (Plant hormone signal transduction pathway) under LD conditions and metabolic processes (Biosynthesis of secondary metabolites) under SD conditions, respectively. Transcriptome analysis revealed *GmFT3a* might promote flowering by regulating reproductive signals and growth regulators. However, the functional complementarity and interactions among multiple *FT* homologues remain to be further verified (Lee et al., 2021). Actually, the interaction and allocation mechanism among the *GmFT* family genes will be further certificated confirmed by evaluation of the phenotypes of the *Gmft3a* mutant and/or the double mutant of different combinations of *Gmft* members in future studies.

The natural variations of *GmFT3a* are related to the high-latitude adaptability of soybean

The adaptation of soybean varieties to different latitudes or daylengths was mainly attributed to the combinations and functional diversifications among the photoperiod-related genes (Liu et al., 2020b; Lin et al., 2021b). The wide distribution of soybean varieties results from rich natural variations and different combinations of genes and QTLs controlling flowering time (Watanabe et al., 2012; Liu et al., 2022). The early maturity of soybeans in Northeast China was largely attributed to local exotic migration and selection against positive alleles causing new recombination, while only few new allele emergence/mutation occurred (Fu et al., 2020). Therefore, allelic variations in flowering time-related genes could be applied as an effective means for breeding (Cao et al., 2016b; Khan et al., 2022). There were natural variations within the *GmFT* family genes that contribute to soybean adaptation to various growing areas (Jiang et al., 2019). Some natural variations in *GmFT2a* caused significantly late flowering, which is suitable for soybean accessions which adapted to low-latitude areas (Li et al., 2021). In contrast, some single-nucleotide polymorphisms (SNPs) in *GmFT5a* result in early flowering, which was preferred in high-latitude areas (Takeshima et al., 2016). For *GmFT3a*, multiple SNP alleles in both promoter and genomic regions suggested that genomic variations and

regulation of gene expression of *GmFT3a* both contributed to the flowering and maturity diversities (Figure 7). Notably, the rare Hap2 of *GmFT3a* showed relatively early maturity and most of these varieties were bred for high-latitude regions in China, suggesting that *GmFT3a*-Hap2 may have contributed to the adaptation of soybean to higher latitude areas. For soybean breeding, it is an effective way to broaden the adaptive region of soybean varieties through recombining and modifying genes in photoperiod pathway, e.g., various combinations of mutations at the *E* loci (*E1*, *E2*, *E3* and *E4*) provide considerable genetic plasticity that contributes to soybean cultivation at diverse latitudes (Jiang et al., 2014; Liu et al., 2020c, 2022); *J* (the major gene for adaptation to short-day and high temperature conditions at low latitudes; Yue et al., 2016; Lu et al., 2017), *GmPRR37* (Wang et al., 2020), the *ft2aft5a* double mutants showed late flowering and produced more pods and seeds under SD conditions (Cai et al., 2020). When using these genes with large impacts on flowering time and maturity, it might be difficult to maintain the original yield and quality traits of elite varieties. However, fine-tuning florigen expression is a promising strategy for high yield (Xu et al., 2021b). In this study, we found that the overexpression of *GmFT3a* promoted flowering by approximately 4.0–8.9 d under LD conditions, suggesting that *GmFT3a* provides an opportunity to lightly modify the flowering time and maturity of soybean varieties, which is helpful to retain yield and agronomic traits of an elite variety with extension of its adaptive regions (Qi et al., 2021).

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 at: https://www.ncbi.nlm.nih.gov/, PRJNA832118.

Author contributions

TH and LY conceived the study. YW and CZ performed the Arabidopsis transformation experiment. LC and WH performed the soybean transformation. JW and LZ recorded the phenotype of transgenic soybean. SS and CW provided and maintained the MG standard varieties. TW and BJ analyzed the re-sequencing data. YC and FC performed the gene expression experiment. SY, XX, and WS analyzed the transcriptomic data. SY and YW drafted the manuscripts. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from National Natural Science Foundation of China (32001566), the Central Public-interest Scientific Institution Basal Research Fund (No. Y2022GH08), and the China Agriculture Research System (CARS-04).

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.2022.929747/full#supplementary-material

References

Beinecke, F. A., Grundmann, L., Wiedmann, D. R., Schmidt, F. J., Caesar, A. S., Zimmermann, M., et al. (2018). The FT/FD-dependent initiation of flowering under long-day conditions in the day-neutral species *Nicotiana tabacum* originates from the facultative short-day ancestor Nicotiana tomentosiformis. *Plant J.* 96, 329–342. doi: 10.1111/tpj.14033

Blümel, M., Dally, N., and Jung, C. (2015). Flowering time regulation in crops-what did we learn from Arabidopsis? *Curr. Opin. Plant Biol.* 32, 121–129. doi: 10.1016/j.copbio.2014.11.023

Cai, Y., Wang, L., Chen, L., Wu, T., Liu, L., Sun, S., et al. (2020). Mutagenesis of *GmFT2a* and *GmFT5a* mediated by CRISPR/Cas9 contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* 18, 298–309. doi: 10.1111/pbi.13199

Cao, K., Cui, L., Zhou, X., Ye, L., Zou, Z., and Deng, S. (2016a). Four tomato FLOWERING LOCUS T-like proteins act antagonistically to regulate floral initiation. *Front. Plant Sci.* 6:1213. doi: 10.3389/fpls.2015.01213

Cao, D., Takeshima, R., Zhao, C., Liu, B., Jun, A., and Kong, F. (2016b). Molecular mechanisms of flowering under long days and stem growth habit in soybean. *J. Exp. Bot.* 68, erw1873–erw1884. doi: 10.1093/jxb/erw394

Chen, L., Cai, Y., Liu, X., Yao, W., Guo, C., Sun, S., et al. (2018). Improvement of soybean *Agrobacterium*-mediated transformation efficiency by adding glutamine and asparagine into the culture media. *Int. J. Mol. Sci.* 19, 3039. doi: 10.3390/ijms19103039

Chen, L., Cai, Y., Qu, M., Wang, L., Sun, H., Jiang, B., et al. (2020). Soybean adaption to high-latitude regions is associated with natural variations of *GmFT2b*, an ortholog of *FLOWERING LOCUS T*. *Plant Cell Environ.* 43, 934–944. doi: 10.1111/pce.13695

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

Cockram, J., Jones, H., Leigh, F. J., O'Sullivan, D., Powell, W., Laurie, D. A., et al. (2007). Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *J. Exp. Bot.* 58, 1231–1244. doi: 10.1093/jxb/erm042

Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q. Z., Searle, L., 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

Fehr, W. R., and Caviness, C. E. (1977). "Stages of soybean development," in *Cooperative Extension Service, Agriculture and Home Economics Experiment Station*. eds. C. E. Donhowe (Ames, IA, USA: Iowa State University of Science and Technology).

Fu, M., Wang, Y., Ren, H., Du, W., Wang, D., Bao, R., et al. (2020). Genetic dynamics of earlier maturity group emergence in south-to-north extension of Northeast China soybeans. *Theor. Appl. Genet.* 133, 1839–1857. doi: 10.1007/s00122-020-03558-4

Hartwig, E. E. (1970). Growth and reproductive characteristics of soybeans (*Glycine max* (L.) Merr.) grown under short-day conditions. *Trop. Sci.* 12, 47–53.

He, Y., Chen, T., and Zeng, X. (2020). Genetic and epigenetic understanding of seasonal timing of flowering. *Plant Commun.* 1, 13. doi: 10.1016/j.xplc.2019.100008

Jia, H., Jiang, B., Wu, C., Hou, W., Sun, S., Yan, H., et al. (2014). Maturity group classification and maturity locus genotyping of early-maturing soybean varieties from high-latitude cold regions. *PLoS One* 9:e94139. doi: 10.1371/journal.pone.0094139

Jiang, B., Nan, H., Gao, Y., Tang, L., Yue, Y., Lu, S., et al. (2014). Allelic combinations of soybean maturity loci *E1*, *E2*, *E3* and *E4* result in diversity of maturity and adaptation to different latitudes. *PLoS One* 9:e106042. doi: 10.1371/journal.pone.0106042

Jiang, B., Zhang, S., Song, W., Abdul, M., Khan, A., Sun, S., et al. (2019). Natural variations of *FT* family genes in soybean varieties covering a wide range of maturity groups. *BMC Genomics* 20, 230. doi: 10.1186/s12864-019-5577-5

Jin, S., Nasim, Z., Susila, H., and Ahn, J. H. (2021). Evolution and functional diversification of *FLOWERING LOCUS T/TERMINAL FLOWER* 1 family genes in plants. *Semin. Cell Dev. Biol.* 109, 20–30. doi: 10.1016/j.semcdb.2020.05.007

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., et al. (1999). Activation tagging of the floral inducer FT. *Science* 286, 1962–1965. doi: 10.1126/science.286.5446.1962

Khan, M. A. A., Zhang, S. W., Emon, R. M., Chen, F., Song, W., Wu, T., et al. (2022). CONSTANS polymorphism modulates flowering time and maturity in soybean. *Front. Plant Sci.* 13:817544. doi: 10.3389/fpls.2022.817544

Kobayashi, Y., and Weigel, D. (2007). Move on up, it's time for change mobile signals controlling photoperiod-dependent flowering. *Genes Dev.* 21, 2371–2384. doi: 10.1101/gad.1589007

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B. M., Watanabe, S., et al. (2010). Two coordinately regulated homologues of *FLOWERING LOCUS T* are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* 154, 1220–1231. doi: 10.1104/pp.110.160796

Lee, S. H., Choi, C. W., Park, K. M., Jung, W. H., Chun, H. J., Baek, D., et al. (2021). Diversification in functions and expressions of soybean *FLOWERING LOCUS T* genes fine-tunes seasonal flowering. *Front. Plant Sci.* 12:613675. doi: 10.3389/fpls.2021.613675

Li, B., and Dewey, C. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinf.* 12, 93–99. doi: 10.1186/1471-2105-12-323

Li, X., Fang, C., Yang, Y., Lv, T., Su, T., Chen, L., et al. (2021). Overcoming the genetic compensation response of soybean florigens to improve adaptation and yield at low latitudes. *Curr. Biol.* 31, 3755–3767.e4. doi: 10.1016/j.cub.2021.06.037

Li, X., Wu, C., Ma, Q., Zhang, S., Li, C., Zhang, X., et al. (2005). Morphology and anatomy of the differentiation of flower bud and the process of flowering reversion in soybean cv. *Zigongdongdou. Acta Agron. Sin.* 31, 1437–1442. doi: 10.3321/j.issn:0496-3490.2005.11.008

Lin, X., Fang, C., Liu, B., and Kong, F. (2021a). Natural variation and artificial selection of photoperiodic flowering genes and their applications in crop adaptation. *aBIOTECH* 2, 156. doi: 10.1007/s42994-021-00039-0

Lin, X., Liu, B., Weller, J. L., Abe, J., and Kong, F. (2021b). Molecular mechanisms for the photoperiodic regulation of flowering in soybean. *J. Integr. Plant Biol.* 63, 981–994. doi: 10.1111/jipb.13021

Liu, L., Gao, L., Zhang, L., Cai, Y., Song, W., Chen, L., et al. (2022). Co-silencing *E1* and its homologs in an extremely late-maturing soybean cultivar confers super-early maturity and adaptation to high-latitude short-season regions. *J. Integr. Agric.* 21, 326–335. doi: 10.1016/S2095-3119(20)63391-3

Liu, X., He, J., Wang, Y., Xing, G., Li, Y., Yang, S., et al. (2020a). Geographic differentiation and phylogeographic relationships among world soybean populations. *Crop J.* 8, 260–272. doi: 10.1016/j.cj.2019.09.010

Liu, W., Jiang, B., Ma, L., Zhang, S., Zhai, H., Xu, X., et al. (2018). Functional diversification of *FLOWERING LOCUS T* homologs in soybean: *GmFT1a* and *GmFT2a*/*5a* have opposite roles in controlling flowering and maturation. *New Phytol.* 217, 1335–1345. doi: 10.1111/nph.14884

Liu, L., Song, W., Wang, L., Sun, X., Qi, Y., Wu, T., et al. (2020c). Allele combinations of maturity genes *E1-E4* affect adaptation of soybean to diverse geographic regions and farming systems in China. *PLoS One* 15:e0235397. doi: 10.1371/journal.pone.0235397

Liu, L., Xuan, L., Jiang, Y., and Yu, H. (2020b). Regulation by *FLOWERING LOCUS T* and *TERMINAL FLOWER 1* in flowering time and plant architecture. *Small Struct.* 2:2000125. doi: 10.1002/sstr.202000125

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean *J* locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49, 773–779. doi: 10.1038/ng.3819

Lv, X., Zeng, X., Hu, H., Chen, L., Zhang, F., Liu, R., et al. (2021). Structural insights into the multivalent binding of the Arabidopsis *FLOWERING LOCUS T* promoter by the CO-NF-Y master transcription factor complex. *Plant Cell* 33, 1182–1195. doi: 10.1093/plcell/koab016

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.2307/1592215

Matias-Hernandez, L., Aguilar-Jaramillo, A. E., Cigliano, R. A., Sanseverino, W., and Pelaz, S. (2016). Flowering and trichome development share hormonal and transcription factor regulation. *J. Exp. Bot.* 67, 1209–1219. doi: 10.1093/jxb/erv534

Mourtzinis, S., and Conley, S. P. (2017). Delineating soybean maturity groups across the United States. *Agron. J.* 109, 1397–1403. doi: 10.2134/agronj2016.10.0581

Na, X., Jian, B., Yao, W., Wu, C., Hou, W., Jiang, B., et al. (2013). Cloning and functional analysis of the flowering gene *GmSOC1*-like, a putative *SUPPRESSOR OF OVEREXPRESSION CO1/AGAMOUS-LIKE 20* (*SOC1/AGL20*) ortholog in soybean. *Plant Cell Rep.* 32, 1219–1229. doi: 10.1007/s00299-013-1419-0

Nakamichi, N. (2015). Adaptation to the local environment by modifications of the photoperiod response in crops. *Plant Cell Physiol.* 56, 594–604. doi: 10.1093/pcp/pcu181

Nan, H., Cao, D., Zhang, D., Li, Y., Lu, S., Tang, L., et al. (2014). *GmFT2a* and *GmFT5a* redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor *GmFDL19* in soybean. *PLoS One* 9:e97669. doi: 10.1371/journal.pone.0097669

Park, S. J., Jiang, K., Tal, L., Yichie, Y., Gar, O., Zamir, D., et al. (2014). Optimization of crop productivity in tomato using induced mutations in the florigen pathway. *Nat. Genet.* 46, 1337–1342. doi: 10.1038/ng.3131

Pin, P. A., Benlloch, R., Bonnet, D., Wremerth-Weich, E., Kraft, T., Gielen, J., et al. (2010). An antagonistic pair of *FT* homologs mediates the control of flowering time in sugar beet. *Science* 330, 1397–1400. doi: 10.1126/science.1197004

Pin, P. A., and Nilsson, O. (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

Qi, X., Jiang, B., Wu, T., Sun, S., Wang, C., Song, W., et al. (2021). Genomic dissection of widely planted soybean cultivars leads to a new breeding strategy of crops in the post-genomic era. *Crop J.* 9, 1079–1087. doi: 10.1016/j.cj.2021.01.001

Qin, Z., Wu, J., Geng, S., Feng, N., Chen, F., Kong, X., et al. (2017). Regulation of *FT* splicing by an endogenous cue in temperate grasses. *Nat. Commun.* 8, 14320. doi: 10.1038/ncomms14320

Ratcliffe, O. J., and Riechmann, J. L. (2002). Arabidopsis transcription factors and the regulation of flowering time: a genomic perspective. *Curr. Issues Mol. Biol.* 296, 285–289. doi: 10.1126/science.296.5566.285

Song, W., Sun, S., Ibrahim, S. E., Xu, Z., Wu, H., Hu, X., et al. (2019). Standard cultivar selection and digital quantification for precise classification of maturity groups in soybean. *Crop Sci.* 59, 1997–2006. doi: 10.2135/cropsci2019.02.0095

Srikanth, A., and Schmid, M. (2011). Regulation of flowering time: all roads lead to Rome. *Cell. Mol. Life Sci.* 68, 2013–2037. doi: 10.1007/s00018-011-0673-y

Su, Q., Chen, L., Cai, Y., Chen, Y., Yuan, S., Li, M., et al. (2022). Functional redundancy of *FLOWERING LOCUS T 3b* in soybean flowering time regulation. *Int. J. Mol. Sci.* 23, 2497. doi: 10.3390/ijms23052497

Sun, H., Jia, Z., Cao, D., Jiang, B., Wu, C., Hou, W., et al. (2011). *GmFT2a*, a soybean homologue of FLOWERING LOCUS T, is involved in flowering transition and maintenance. *PLoS One* 6:e29238. doi: 10.1371/journal.pone.0029238

Takeshima, R., Hayashi, T., Zhu, J., Zhao, C., Xu, M., Yamaguchi, N., et al. (2016). A soybean quantitative trait locus that promotes FLOWERING under long days is identified as *FT5a*, a *FLOWERING LOCUS T* ortholog. *J. Exp. Bot.* 67, 5247–5258. doi: 10.1093/jxb/erw283

Takeshima, R., Nan, H., Harigai, K., Dong, L., and Abe, J. (2019). Functional divergence between soybean *FLOWERING LOCUS T* orthologues, *FT2a*. *J. Exp. Bot.* 70, 3941–3953. doi: 10.1093/jxb/erz199

Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., and Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. *Science* 316, 1033–1036. doi: 10.1126/science.1141753

Thomas, B., and Vince-Prue, D. (1997). *Photoperiodism in Plants*. USA: Academic Press, Inc.

Wang, L., Sun, S., Wu, T., Liu, L., Sun, X., Cai, Y., et al. (2020). Natural variation and CRISPR/Cas9-mediated mutation in *GmPRR37* affect photoperiodic flowering and contribute to regional adaptation of soybean. *Plant Biotechnol. J.* 18, 1869–1881. doi: 10.1111/pbi.13346

Wang, Z., Zhou, Z., Liu, Y., Liu, T., Li, Q., Ji, Y., et al. (2015). Functional evolution of phosphatidylethanolamine binding proteins in soybean and Arabidopsis. *Plant Cell* 27, 323–336. doi: 10.1105/tpc.114.135103

Watanabe, S., Harada, K., and Abe, J. (2012). Genetic and molecular bases of photoperiod responses of flowering in soybean. *Breed. Sci.* 61, 531–543. doi: 10.1270/jsbbs.61.531

Wigge, P. (2011). FT, A mobile developmental signal in plants. *Curr. Biol.* 21, R374–R378. doi: 10.1016/j.cub.2011.03.038

Wu, T., Li, J., Wu, C., Sun, S., Mao, T., Jiang, B., et al. (2015). Analysis of the independent-and interactive-photo-thermal effects on soybean flowering. *J. Integr. Agric.* 14, 622–632. doi: 10.1016/S2095-3119(14)60856-X

Wu, C., Ma, Q., Yam, K. M., Cheung, M. Y., Xu, Y., Han, T., et al. (2006). In situ expression of the *GmNMH7* gene is photoperiod-dependent in a unique soybean (*Glycine max* [L.] Merr.) flowering reversion system. *Planta* 223, 725–735. doi: 10.1007/s00425-005-0130-y

Xu, X., Zhang, L., Cao, X., Liu, L., Jiang, B., Zhang, C., et al. (2021a). Cotyledons facilitate the adaptation of early maturing soybean varieties to high latitude long day environments. *Plant Cell Environ.* 44, 2551–2564. doi: 10.1111/pce.14120

Xu, K., Zhang, X. M., Chen, H., Zhang, C., Zhu, J., Cheng, Z., et al. (2021b). Fine-tuning florigen increases field yield through improving photosynthesis in soybean. *Front. Plant Sci.* 12:710754. doi: 10.3389/fpls.2021.710754

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11, R14. doi: 10.1186/gb-2010-11-2-r14

Yue, Y., Liu, N., Jiang, B., Li, M., Wang, H., Jiang, Z., et al. (2016). A single nucleotide deletion in *J* encoding *GmELF3* confers long juvenility and is associated with adaption of tropic soybean. *Mol. Plant* 10, 656–658. doi: 10.1016/j.molp.2016.12.004

Zeevaart, J. A. (2008). Leaf-produced floral signals. *Curr. Opin. Plant Biol.* 11, 541–547. doi: 10.1016/j.pbi.2008.06.009

Zhai, H., Lü, S., Liang, S., Wu, H., Zhang, X., Liu, B., et al. (2014). *GmFT4*, a homologue of FLOWERING LOCUS T, is positively regulated by *E1* and functions as a flowering repressor in soybean. *PLoS One* 9:e89030. doi: 10.1371/journal.pone.0089030

Zhang, L., Liu, W., Mesfin, T., Xu, X., Qi, Y., Sapey, E., et al. (2020). Principles and practices of the photo-thermal adaptability improvement in soybean. *J. Integr. Agric.* 19, 295–310. doi: 10.1016/S2095-3119(19)62850-9

Zhang, S., Singh, M. B., and Bhalla, P. L. (2021). Molecular characterization of a soybean *FT* homologue, *GMFT7*. *Sci. Rep.* 11, 3651. doi: 10.1038/s41598-021-83305-x

Zhang, T., Wu, T., Wang, L., Jiang, B., Zhen, C., Yuan, S., et al. (2019). A combined linkage and GWAS analysis identifies QTLs linked to soybean seed protein and oil content. *Int. J. Mol. Sci.* 20, 5915. doi: 10.3390/ijms20235915

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 14 October 2022
DOI 10.3389/fpls.2022.955623

OPEN ACCESS

EDITED BY
Liang Wu,
Zhejiang University, China

REVIEWED BY
Koen Geuten,
KU Leuven, Belgium
Fa Cui,
Ludong University, China
Aili Li,
Chinese Academy of Agricultural Sciences (CAAS), China

*CORRESPONDENCE
Ben Trevaskis
ben.trevaskis@csiro.au

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 29 May 2022
ACCEPTED 20 September 2022
PUBLISHED 14 October 2022

CITATION
Trevaskis B, Harris FAJ, Bovill WD, Rattey AR, Khoo KHP, Boden SA and Hyles J (2022) Advancing understanding of oat phenology for crop adaptation.
Front. Plant Sci. 13:955623.
doi: 10.3389/fpls.2022.955623

Advancing understanding of oat phenology for crop adaptation

Ben Trevaskis[1]*, Felicity A. J. Harris[2,3], William D. Bovill[1], Allan R. Rattey[4], Kelvin H. P. Khoo[5], Scott A. Boden[5] and Jessica Hyles[1]

[1]Commonwealth Scientific and Industrial Research Organisation, Agriculture and Food Business Unit, Black Mountain Science and Innovation Park, Canberra, ACT, Australia, [2]Department of Primary Industries, Pine Gully Road, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW, Australia, [3]School of Agricultural, Environmental and Veterinary Sciences, Charles Sturt University, Wagga Wagga, NSW, Australia, [4]Intergrain, Perth, WA, Australia, [5]School of Agriculture, Food & Wine, Faculty of Sciences, Waite Research Institute, University of Adelaide, Urrbrae, Adelaide, SA, Australia

Oat *(Avena sativa)* is an annual cereal grown for forage, fodder and grain. Seasonal flowering behaviour, or phenology, is a key contributor to the success of oat as a crop. As a species, oat is a vernalization-responsive long-day plant that flowers after winter as days lengthen in spring. Variation in both vernalization and daylength requirements broadens adaptation of oat and has been used to breed modern cultivars with seasonal flowering behaviours suited to different regions, sowing dates and farming practices. This review examines the importance of variation in oat phenology for crop adaptation. Strategies to advance understanding of the genetic basis of oat phenology are then outlined. These include the potential to transfer knowledge from related temperate cereals, particularly wheat (*Triticum aestivum*) and barley (*Hordeum vulgare*), to provide insights into the potential molecular basis of variation in oat phenology. Approaches that use emerging genomic resources to directly investigate the molecular basis of oat phenology are also described, including application of high-resolution genome-wide diversity surveys to map genes linked to variation in flowering behaviour. The need to resolve the contribution of individual phenology genes to crop performance by developing oat genetic resources, such as near-isogenic lines, is emphasised. Finally, ways that deeper knowledge of oat phenology can be applied to breed improved varieties and to inform on-farm decision-making are outlined.

KEYWORDS

oat, flowering, seasons, vernalization, photoperiod

Oat

Oat (*Avena sativa*) was potentially a wild food for Paleolithic hunter-gatherers (Lippi et al., 2015), before entering agriculture as a weed of wheat and barley during the Neolithic period (Murphy and Hoffman, 1992). As a weed, oat can dominate fields of wheat or barley and by 2000 years ago it was itself being grown as a crop (Harlan, 1982; Murphy and Hoffman, 1992). Factors that drove the acceptance of oat into early European farming systems included the ability to perform well in colder climates and marginal areas, with less inputs, together with end-use versatility that contributed to overall farm resilience (Moore-Colyer, 1995). The nutritional quality of oat grains, or groats, was also a key factor. Oat grains have high lipid content, contain lysine-rich protein and soluble fibre, particularly β-glucan (Cuddeford, 1995; Zwer, 2017). Oat grains are an ideal feed for horses, so there was strong demand for oats throughout the period when horses were a major contributor to transport and industry (Murphy and Hoffman, 1992). The same grain quality parameters are now driving interest in oats as a healthy food and as a non-animal protein source.

The importance of phenology to oat adaption

Oat is an important crop in Australia, Canada, China, Europe, North and South America. Variation in the seasonal timing of life cycle events (phenology), particularly the timing of flowering and grain production, enables adjustment of the oat life cycle to suit local constraints. This variation is critical for the cultivation of modern oats across such a broad geographical range. For example, cultivation of oat in Canada is typically limited to spring and summer, thereby avoiding harsh winters. This contrasts with regions such as the United Kingdom where crops can be sown in autumn, then over-winter before producing grain in spring and summer. Other climate factors that can define the timing and duration of the growing season for oat include seasonal rainfall patterns and extreme summer heat.

The history of the Australian grains industry provides an example of the importance of variation in phenology for geographical adaptation. Oat came to Australia with Europeans and was cultivated in the Sydney area as early as 1791 (Collins and King, 1798). It was then slow to expand as a crop due to a lack of varieties suited to Australian growing conditions. This was because the oat growing season in south-eastern Australia is constrained by the timing of autumn rains, which dictate sowing dates, combined with the need to flower in the "optimal flowering period" to avoid frost, heat and water-limitation disrupting flowering and grain development (Tashiro and Wardlaw, 1990; Dolferus et al., 2011; Flohr et al., 2017). The European oats first introduced to Australia were ill-suited to the local growing conditions because they were slow to mature and flower. This remained the case until the early 20th century when John Pridham, a cereal breeder trained by the pioneer wheat breeder William Farrer, began to breed locally adapted oat cultivars (Mengersen, 1960; Fitzsimmons, 1988). Pridham selected oats with earlier flowering, with his most successful variety being Belar (Fitzsimmons, 1988). Breeding of oats for Australian growing conditions continued with Pridham's successors and this facilitated the expansion of the Australian oat industry (Mengersen, 1960). Now, on average, 1.4 million tons of milling oats are produced annually, with Western Australia and New South Wales being the major growing regions (ABARES, 2021). Additionally, around 0.7 million tons of oaten hay are exported annually, and more is used domestically as fodder (ABARES, 2021). Oat is also grown as a forage crop across large areas of New South Wales and Queensland.

Similar examples of how phenology adapts oat to local seasonal constraints can be found for other regions globally, though the specific of constraints differ to those in Australia (harsh winters in Canadian prairies, for example). In addition, as outlined in subsequent sections, phenology influences yield component traits, such as grain number, *via* determining pattens of plant growth and development, and so influences crop yield potential. The overall relationship between phenology and crop yield, through adaptation and yield potential, is not limited to Australian context. For example, phenology is a driver of yield in North American and European climates (Yan et al., 2007; Howarth et al., 2021).

From the perspective of overall farming systems, phenology also allows adaption to specific farm management practices. For example, selection for earlier flowering can allow double cropping in some farming systems (Islam et al., 2010). There is also evidence that in some regions breeders have selected for earlier maturity during the transition to mechanisation, which allowed sowing of large areas to occur within a narrower seasonal window (Grau Nersting et al., 2006). Similarly, phenology is an important component of optimisation of oat for different end uses. For example, the timing of maturation influences quality of milling oats (Howarth et al., 2021). Interactions between phenology and seasonal conditions also influence forage oat quality (Coblentz et al., 2012; Ki-Seung et al., 2014).

Phenology is likely to be of renewed global importance in an era of warming climates and increased climate variability. Re-optimisation of phenology might be required in some production zones. This could include adjustment of phenology to suit shifting sowing or harvest dates, and/or adjustment of phenology to compensate for more rapid growth and development driven by faster accumulation of growing degree days (Hunt et al., 2019; Peltonen-Sainto and Juahianen, 2020).

In summary, phenology enables adaption of oat to different climates, is a major determinant of crop yields, and can enable

breeding of cultivars for different farming systems or end-uses. Subsequent sections review current understanding of oat phenology.

Seasonal flowering behaviour of oat

The time of year when oat flowers and produces grain is determined largely by genotype-dependent responses to seasonal temperature and daylength (photoperiod) cues. As a species, cultivated oat can be described as a "vernalization-responsive long-day plant", which will flower after exposure to winter cold (i.e., vernalization) as days lengthen in spring (Figure 1) (Borodin, 1934; Wiggans and Frey, 1955). This archetypal combination of seasonal flowering-responses is found in "winter" oats, which are well-suited to situations where crops are sown autumn and then overwinter before flowering in spring (Sampson and Burrows, 1972; Sorrells and Simmons, 1992). There is variation in both vernalization and daylength requirements amongst oat accessions/cultivars (Sampson and Burrows, 1972: King and Bacon, 1992). This variation has been harnessed by plant breeders to selectively modify oat phenology to suit diverse environments and different management practices, as outlined above.

Winter oats can require up to 50 days at low temperatures (below 10°C) to flower rapidly (Sorrells and Simmons, 1992). The response to cold is quantitative, such that increasing the duration of vernalization typically accelerates flowering to greater extents until the vernalization response is saturated (Sampson and Burrows, 1972). Vernalization accelerates the transition to flowering, the point when the shoot apex switches from vegetative to reproductive development (Bell, 1936; Bonnett, 1966; Aitken, 1974). A clear developmental indicator of the vernalization response, aside from earlier flowering, is reduction in the number of leaves produced on the main stem (final leaf number), since a vernalized plant spends less time in

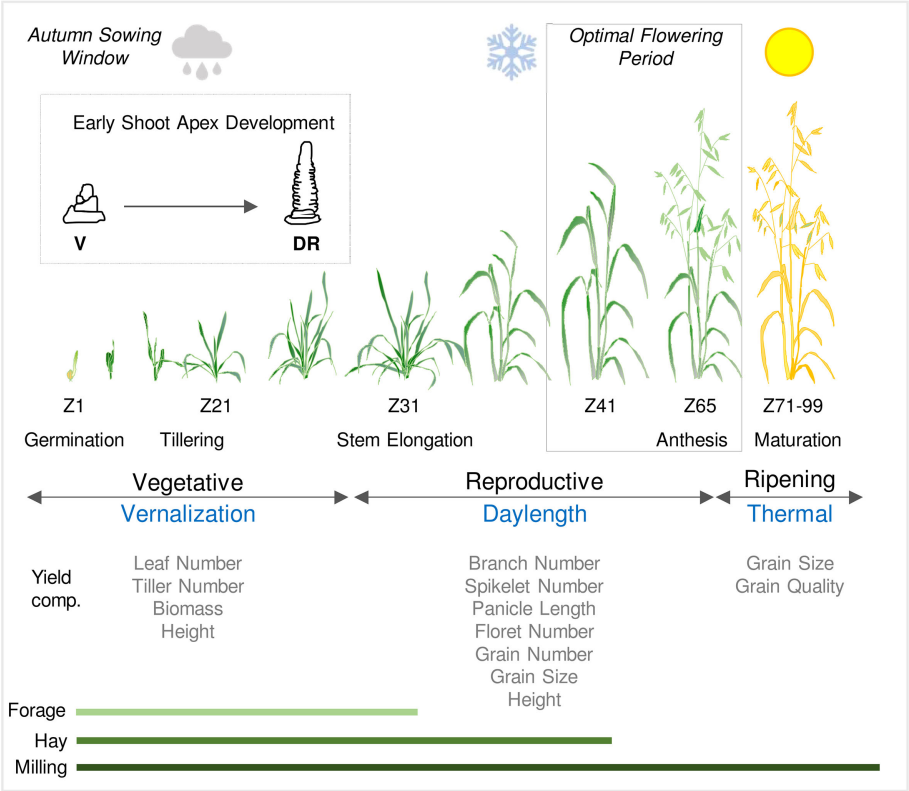


FIGURE 1
Schematic representation of the life cycle of oat in an Australian context. Sowing and establishment of crops is timed according to water availability in autumn (cloud symbol) and then plants experience cold temperatures through late autumn and winter (vernalization). Flowering and grain production occur during a optimal flowering period in spring, after the risk of frost subsides (snowflake symbol) and before the onset of heat and water limitation (sun symbol). Developmental stages are referenced by Zadoks scale (Zadoks et al., 1974). Inset shows early shoot apex development and the first visible sign of the transition from vegetative (V) to reproductive development; the double ridge stage (DR). The relationships between phases of development (vegetative and reproductive) and ripening are shown relative to the seasonal climate cues that influence the duration of each these phases. The impact of phase durations on physiological and yield related traits are indicated together with the relevance of growth phases to different end uses (forage, hay or milling).

the vegetative phase, generating fewer phytomers and less leaves over the course of the plant life cycle (Aitken, 1974). A saturating vernalization treatment is one that generates the minimum final leaf number (Aitken, 1974). Whereas winter oats require vernalization to flower, spring oats generally show reduced or no requirement for vernalization (Borodin, 1934; Sampson and Burrows, 1972: King and Bacon, 1992). This allows spring oats to be sown when vernalization might not occur, after winter or in regions where winters are mild.

Daylengths exceeding 12 hours (long days) typically accelerate flowering of oat. The extent to which flowering is accelerated increases with longer daylengths, with the maximal response occurring when daylength exceeds 18 hours (Sampson and Burrows, 1972; Sorrells and Simmons, 1992). Daylength can influence the duration of the vegetative growth phase but, under typical field conditions, increasing daylengths during spring typically coincide with reproductive development and so have more impact on the duration of inflorescence development (Jenkins, 1973; Sorrells and Simmons, 1992) (Figure 1). While flowering of oat is usually delayed when plants are grown in daylengths shorter than 12 hours, some varieties flower rapidly irrespective of daylength (Burrows, 1984). There are examples of oats that flower rapidly irrespective of both vernalization and daylength (Aitken, 1974).

Cereal phenology genes

For the reasons outlined above, selecting optimal phenology is a key goal for oat breeding programs, so understanding the genetic basis for variation in oat phenology can contribute to future breeding strategies. The molecular basis of variation in oat phenology has not been resolved. There is, however, detailed knowledge of the gene sequences that underlie variation in the vernalization and daylength requirements of other temperate cereals, particularly wheat and barley (see Fjellheim et al., 2014; Hyles et al., 2020). The *Avena* genus is part of the same subfamily of grasses as wheat and barley (Pooideaea) and all are members of the "Core Pooid" clade. This close evolutionary relationship, combined with the similar flowering physiology, suggests that there are good prospects to transfer knowledge of the molecular basis of seasonal flowering from wheat and barley to oat.

Vernalization

The key gene controlling vernalization-induced flowering of wheat and barley is the MADS box transcription factor, *VERNALIZATION1* (*VRN1*) (Danyluk et al., 2003; Trevaskis et al., 2003; Yan et al., 2003). *VRN1* promotes the transition to reproductive development but is transcribed at low levels in plants that have not been vernalized. Exposing plants to prolonged cold activates transcription of *VRN1* and this subsequently accelerates flowering when plants are grown at warmer temperatures (Danyluk et al., 2003; Trevaskis et al., 2003; Yan et al., 2003; Chen and Dubcovsky, 2012). Activation of *VRN1* by cold is quantitative such that *VRN1* transcript levels increase more with longer cold treatments, accelerating flowering to greater extents (Danyluk et al., 2003; Trevaskis et al., 2003; Yan et al., 2003; von Zitzewitz et al., 2005; Sasani et al., 2009). Spring wheats that flower without vernalization typically carry alleles of *VRN1* that are transcribed without cold and so bypass the requirement for vernalization (Pugsley, 1971; Yan et al., 2003; Yan et al., 2004a; Fu et al., 2005; Eagles et al., 2009; Eagles et al., 2011; Díaz et al., 2012; Oliver et al., 2013). These alleles have mutations in the promoter or first intron that are suggested to trigger elevated transcriptional activity (Yan et al., 2003; Yan et al., 2004a). *VRN1* copy number variation and amino acid substitutions have also been linked to variation in vernalization requirement of wheat (Chen et al., 2009; Díaz et al., 2012; Dixon et al., 2019).

Daylength

The mechanisms underlying the daylength flowering response were first resolved in *Arabidopsis thaliana* (Arabidopsis) where acceleration of flowering by long days is mediated by *FLOWERING LOCUS T* (*FT*) (Kardailsky et al., 1999; Kobayashi et al., 1999; Corbesier et al., 2007). *FT* encodes a small protein, often described as a florigen, that is expressed in leaves in long days and then transported to the shoot apex where it triggers flowering (Corbesier et al., 2007; Tamaki et al., 2007). Activation of *FT* expression in leaves is activated by a molecular network that includes phytochromes, which perceive external light cues, and the circadian oscillator, which gives rise to rhythmic day-night gene expression patterns that mediate daylength (or photoperiod) responses (Suarez-Lopez et al., 2001; Valverde et al., 2004). The photoperiod response pathway first elucidated in Arabidopsis seems to be widely conserved in plants, including wheat and barley, where transcription of a *FT*-like gene (*FT1*, also known as *VERNALIZATION3, VRN3*) is activated in the leaves by long days (Turner et al., 2005; Yan et al., 2006).

Mutations in genes that mediate circadian oscillator function and output have been linked to variation for daylength sensitivity. For example, the main gene that determines daylength sensitivity of wheat and barley is *PHOTOPERIOD1 (PPD1)* (Turner et al., 2005). *PPD1* is related to *PSEUDORESPONSE REGULATOR* genes that contribute to circadian oscillator function (Turner et al., 2005; Beales et al., 2007). In photoperiod sensitive wheats, which require long days to flower, *PPD1* is expressed with an oscillating expression pattern through day-night cycles and mediates activation of *FT* expression when days are long or increasing in length (i.e.,

spring and summer) (Beales et al., 2007). Mutations that alter *PPD1* activity are a common basis for variation in the photoperiod sensitivity of wheat and barley. For example, a deletion in the promoter of the D genome of *PPD1* is associated with high *PPD1* expression throughout day-night cycles and, in turn, with elevated *FT1* expression and earlier flowering in short days (Beales et al., 2007; Gauley and Boden, 2021). Conversely, coding sequence mutations in the barley *PPD1* gene that likely inhibit function of the PPD1 protein are linked to delayed flowering under long days (Turner et al., 2005). Copy number variation for the *PPD1* gene has also been linked to variation in photoperiod sensitivity of wheat (Díaz et al., 2012). Other genes that function in the circadian oscillator influence phenology of wheat and barley, including *LUX/ARRYTHMO* and *EARLY FLOWERING 3* (Mizuno et al., 2012; Faure et al., 2012; Zakhrabekova et al., 2012; Campoli et al., 2013; Boden et al., 2014; Gawroński et al., 2014; Zikhali et al., 2016). Additionally, *PHYTOCHROME C (PHYC)* plays a major role in daylength perception in cereals and can influence the photoperiod flowering response (Chen et al., 2014). Similar to *PPD1*, all of these genes seem to influence flowering behaviour, at least in part, by influencing *FT1* expression (Figure 2).

Integration of vernalization and daylength pathways

Vernalization is normally a pre-requisite for long-day induction of *FT1* (Hemming et al., 2008). This is mediated by the *VERNALIZATION2 (VRN2)* gene, which limits expression of *FT1* in long days to repress flowering prior to vernalization (Yan et al., 2004b; Trevaskis et al., 2007; Hemming et al., 2008). Loss of *VRN2* function (e.g., through gene deletion) allows long-day induction of *FT1* expression without plants experiencing prolonged cold and thereby bypasses the vernalization requirement (Yan et al., 2004b; Hemming et al., 2008). This occurs mainly in barley and diploid einkorn wheat but is unlikely to be common in tetraploid durum or hexaploid bread wheats, where gene redundancy reduces the likelihood of complete loss of *VRN2* function. Alleles of *FT1* that are expressed at high basal

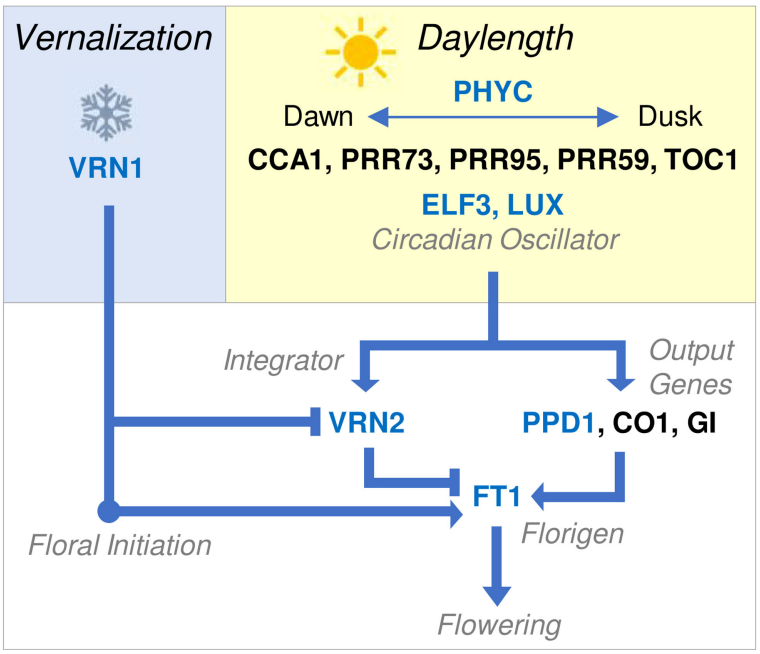


FIGURE 2
Overview of molecular genetic network controlling the seasonal flowering behaviour of temperate cereals. Vernalization (blue box) activates *VRN1*. This is sufficient to trigger the transition to reproductive development or floral initiation, as indicated by the double ridge stage of shoot apex development. Additionally, *VRN1* represses *VRN2* and activates *FT1*, allowing the daylength flowering response to occur after vernalization. Daylength perception (yellow box) is mediated by the circadian oscillator, which generates internal biological rhythms that allow perception of external photoperiod. The circadian oscillator regulates output genes (white box) that activate *FT1*, the key trigger for daylength-induced flowering also known as florigen. *VRN2* is also daylength responsive and blocks long-day induction of *FT1*, until *VRN2* is itself repressed by *VRN1*. *FT1* accelerates reproductive development and stem elongation, leading to flowering. Arrows indicate activation of a target gene, lines ending in bars indicate repression. Blue text indicates that a gene has been linked to (or associated with) natural variation for seasonal flowering behaviour. Gene names are shown without italics for clarity and blue text indicates genes known to mediate natural variation in phenology of other cereals.

levels can bypass *VRN2* and trigger rapid flowering irrespective of vernalization status or daylength (Yan et al., 2006). The precise genetic basis for these alleles has not been resolved; copy number, local rearrangements and insertions/deletions have all been associated with 'early' alleles of *FT1* (Yan et al., 2006; Loscos et al., 2014; Nitcher et al., 2014).

The relationship between phenology genes, development and yield component traits of cereals

Phenology influences other aspects of oat biology in addition to determining when flowering and grain production occur. Vernalization can alter tiller number and there is a strong relationship between the duration of vegetative growth phase and the capacity to survive freezing winter temperatures, as is the case for other temperate cereals (Sorrells and Simmons, 1992; Wooten et al., 2009). By influencing the duration of inflorescence development, the daylength flowering response influences inflorescence structure, such that longer daylengths reduce the duration of inflorescence development and decrease the number of spikelets produced by the oat panicle (Finkner et al., 1973; Peltonen-Sainto, 1994). Genes that determine the duration of vegetative or reproductive development can influence inflorescence node and spikelet number of wheat, barley and rice (*Oryza sativa*). Examples include the *PPD1* and *FT1* genes of wheat and also the *HEADING DATE1* and *EARLY HEADING DATE 1* genes of rice (Endo-Higashi and Izawa, 2011; Boden et al., 2015; Dixon et al., 2018; Finnegan et al., 2018; Brassac et al., 2021). The common theme from these examples is that genes that reduce the duration of reproductive development can decrease the number of inflorescence nodes produced and thus the number of spikelets and florets produced by each plant. This can in turn impact the total grain number per unit area and thereby influence yield. This parallels observations that vernalization and daylength influence tiller or spikelet number of oat (see above). So, in addition to playing a central role in adaptation, oat phenology genes are likely to influence other traits that underpin grain yield, as is the case for other temperate cereals and rice (see Trevaskis, 2018) (Figure 1).

Genes that drive variation in oat phenology

A role for *VRN1*-like genes in mediating vernalization-induced flowering appears to be broadly conserved in Pooid grasses (Ream et al., 2014; McKeown et al., 2016; Feng et al., 2017; Woods et al., 2017). This suggests that *VRN1*-like genes are likely to play a role in vernalization-induced flowering of oat, a member of the core Pooid grasses. Direct support for this idea comes from observations that an oat *VRN1* orthologue is induced by vernalization in winter cultivars but is expressed at high levels irrespective of vernalization status in spring cultivars (Preston and Kellogg, 2008). Additionally, analysis of *FT1* expression patterns across the Pooid clade suggests that long-day induction of *FT1* was also a feature of the common ancestor of this group of grasses (Ream et al., 2014; McKeown et al., 2016). If the molecular networks controlling the seasonal flowering behaviour of oat are similar to those of wheat and barley, a key question then becomes; does variation for vernalization requirement or photoperiod sensitivity map to oat homologues of *VRN1* or *FT*?

Several studies describe genetic mapping of oat loci that contribute to variation in the timing of flowering (Wight et al., 1994; Holland et al., 1997; Holland et al., 2002; Portyanko et al., 2005; Locatelli et al., 2006; Wooten et al., 2009; Nava et al., 2012; Herrmann et al., 2014; Esvelt Klos et al., 2016; Sunstrum et al., 2019). These studies used a range of mapping populations, including recombinant inbred lines and diversity panels, which were genotyped using different marker systems. Phenotyping was conducted in diverse conditions including controlled environments, with different vernalization or daylength treatments, and at a range of field locations with different sowing dates across different years. Many of these studies linked phenological variation to chromosomal locations where *VRN1* or *FT1* are likely to be located according to cross-species comparisons of genetic maps. For example, Holland et al. (2002) suggested the marker BCD808b, which is linked to variation in vernalization responsiveness, is located near a copy of the *VRN1* gene. Another study amplified *VRN1* and *FT1* gene sequences directly from oat, identified nucleotide polymorphisms in these sequences, and then used these to assign potential locations for these genes on genetic maps (Nava et al., 2012). This study also concluded that *VRN1* is linked to variation in oat phenology.

The assembly of hexaploid (AACCDD) oat reference genome sequences (Maughan et al., 2019; Kamal et al., 2022; Peng et al., 2022) now allows direct comparisons between the physical position of candidate phenology genes and the locations of genetic markers linked to variation in flowering behaviour. These comparisons support the idea that *VRN1* and *FT1* are likely to underlie variation in oat phenology, with *VRN1* homeoalleles on chromosomes 4D, 7A and 7D, and *FT1* homeoallele on 7A all linked to phenological variation (Tinker et al., 2022). *FT1* is also potentially linked to the *DAYLENGTH INSENSITIVITY1 (Di1)* gene, on chromosome 7D, which reduces daylength sensitivity by accelerating flowering in short days and has been used to breed rapid cycling oats for Canada and Brazil (Burrows, 1984; Locatelli et al., 2006; Tinker et al., 2022).

It should be noted that there are other candidate genes within the chromosomal regions of interest outlined above. For example, *PHYC* and an oat homologue of *HEADING DATE 6 (HD6)* (Takahashi et al., 2001) are co-located with *VRN1*. This is

consistent with the suggestion of Sunstrum et al. (2019) that the genetic interval where *VRN1* is located might combine effects from multiple genes, including genes that influence flowering in a vernalization-independent manner. Alternative candidate genes have also been suggested for the region linked to *Di1*, including a *CONSTANS*-like gene. (Chaffin et al., 2016; Sunstrum et al., 2019; Tinker et al., 2022). This region is difficult to resolve because marker-trait relationships are inconsistent across populations and there is also evidence for reduced recombination rates (Chaffin et al., 2016; Sunstrum et al., 2019; Tinker et al., 2022). Nevertheless, *VRN1* and *FT1* genes are promising candidates for more detailed analyses, while *PHYC* and *HD6* also warrant further characterisation.

In summary, variation in oat phenology has been mapped to chromosomal regions containing genes related to those that control vernalization and daylength sensitivity in other cereals. Detailed understanding into the nature of the variation in these genes, and how this affects gene function and overall flowering behaviour, is not yet available.

Advancing understanding of oat phenology for adaptation to Australian farming systems

A deeper understanding of the role that oat phenology genes play in adaptation can contribute to the future success of Australian oats through breeding and by contributing to on farm management decision-making. Additionally, there is potential to use phenology genes to optimise plant architecture and thereby increase crop yield.

A first step towards developing an increased understanding of the genetic basis of oat phenology is to identify "major genes" that control vernalization requirement and photoperiod sensitivity. Current knowledge allows the genetic intervals containing these genes to be identified (see above). Future studies will focus on validating and further characterising "causal genes", as has been achieved in wheat and barley. A number of high-resolution mapping approaches can be used to achieve this aim, but ultimately a test of gene function, such as mutagenesis or gene editing, will be required to verify candidate gene functions (e.g., Boden et al., 2015). Additionally, as has been the case in wheat and barley, detailed analysis of gene expression patterns in diverse oat varieties, in response to different vernalization or photoperiod treatments, can provide further insights into the potential roles of candidate phenology genes.

Already there is scope for surveys of genetic diversity using high-throughput sequencing, either of whole genomes or by targeting candidate genes in the regions already identified by genetic mapping and cross-species comparisons. Sequencing of oat pangenomes (https://wheat.pw.usda.gov/GG3/PanOat) will further accelerate this approach. The collected pedigrees of oat varieties, which have been recorded and curated by breeders over many decades (Tinker and Deyl, 2005), are a valuable resource for surveying genetic diversity of major phenology genes. Access to pedigrees allows diversity surveys to be targeted to accessions and cultivars that represent key lines from the history of oat breeding, including key founders and breeding parents spanning the entire period from development of early modern varieties to current elite cultivars (see Eagles et al., 2009). Understanding the genetic basis of phenology and adaptation of oats that were bred for different regions is a logical approach to understand adaptation to local climate and farming practices.

An example of an oat pedigree is presented for Mitika, a South Australian milling oat that was registered in 2003 (Figure 3). This example highlights how pedigrees can be used to identify the donors of key traits into a local breeding program, including breeding line OT207 for the *DWARF6 (DW6)* reduced height gene, the Canadian cultivars Terra for the *NAKED1* gene for hull-less grain and Dumont for crown rust (*Puccinia coronata*) resistance genes *Pc38* and *Pc39* (Brown et al., 1980; McKenzie et al., 1984; Ubert et al., 2017). The pedigree of Mitika also shows extensive mixing of Australian and North American germplasm. Also noteworthy is that all ancestors shown are described as spring types. This suggests that there is potential to introduce more variation for vernalization requirement into Australian milling oats to suit specific environments or farming systems opportunities, such as earlier sowing and dual-purpose crops (graze and grain).

Knowledge of diversity in the major phenology genes can be used in breeding to facilitate choice of parents with compatible phenology for crossing and/or to select progeny of crosses, using molecular markers. Based on knowledge of wheat and barley there is potential for multiple functional alleles of major phenology genes (i.e., many variants of each gene) that have different impacts on phenology and/or yield (Eagles et al., 2009; Hemming et al., 2009; Eagles et al., 2010; Eagles et al., 2011; Díaz et al., 2012). Molecular markers should ideally have the capacity to resolve different functional haplotypes or alleles (Eagles et al., 2010; Eagles et al., 2011). Design of such markers requires high-resolution assays of genetic variation at key loci, using gene resequencing or high-density genome-wide Single Nucleotide Polymorphism (SNP) assays, for example. There then needs to be functional understanding of how different haplotypes influence phenology. Development of near-isogenic lines, by recurrent backcrossing, is a reliable way to contrast the effects of different alleles on phenology; different haplotypes identified by diversity surveys can be introgressed into a common genetic background for phenotyping in different controlled and field conditions (e.g., Hunt et al., 2019). Ideally a successful modern oat that performs well in target environments would be chosen as the recurrent parent that provides the genetic background for near-isogenic lines (e.g., Mitika or Bannister for Australia). Since oat is transformable there is also potential to develop and test

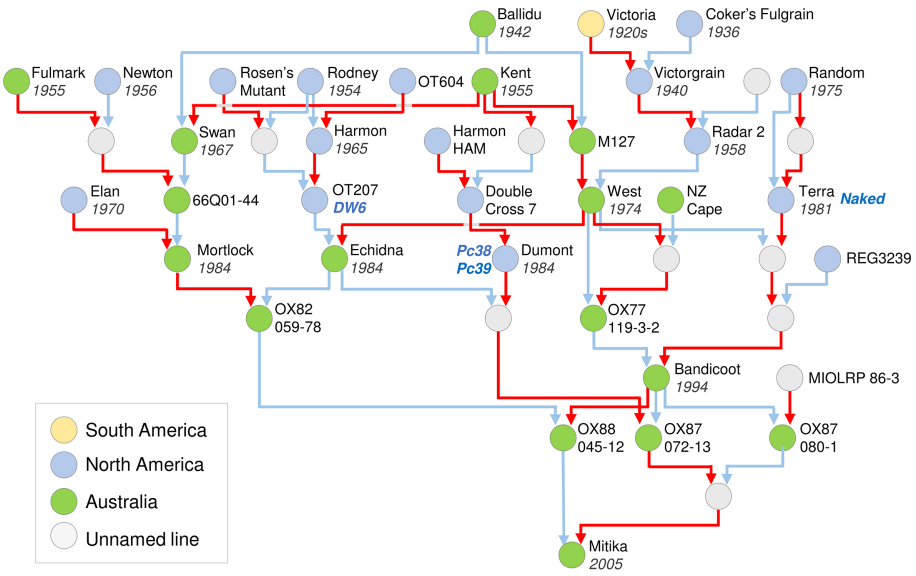


FIGURE 3
Pedigree of an elite Australian spring oat cultivar, cv. Mitika. A simplified pedigree of cultivar Mitika, modified from a Helium output (Shaw et al., 2014) generated from the pedigrees of over 1000 international accessions spanning diversity relevant to the history of Australian oat breeding. Pedigrees are shown as far back as early 20th century oats from Australian and North America. For ease of representation, multiple rounds of backcrossing are not shown. Pedigrees were obtained from the "Pedigrees of Oat Lines" POOL database (Tinker and Deyl, 2005), from Fitzsimmons et al. (1983) and directly from oat breeders (Dr Pamela Zwer and Dr Bruce Winter, personal communication). The year of release is indicated for cultivars. Red line indicates maternal parent connection and blue indicates paternal parent (pollen donor). Colours of circles indicate region of origin for released lines. Grey circles indicate unnamed intermediate lines used in crossing.

novel variation in phenology using gene editing (Gasparis et al., 2008; Matres et al., 2021). Ideally this will target modern elite cultivars that are most relevant to the grains industry.

If the impact of different alleles of the major phenology genes can be resolved then a key question becomes; how can detailed knowledge of individual genes derived from reductionist approaches be re-integrated to understand, predict or optimise overall crop performance? This is a complex challenge that needs to consider both gene-gene and gene-environment interactions. There is also a "sparse data" challenge because detailed phenological observations will only ever be available for specific genotypes grown at particular field sites or sowing dates, representing only a subset of all the scenarios that potentially exist. One effective approach to address these challenges is to integrate genetic knowledge into physiology-based predictive models, such as Agricultural Production Systems Simulation (APSIM, Keating et al., 2003; Zheng et al., 2013). Predictive tools like APSIM can use genetic knowledge to help inform on-farm decision-making. For example, genotypes for vernalization and photoperiod genes can be incorporated into simulation models that predict the date of flowering for specific cultivars, from different sowing dates. Combined with climate data, these predictions can be used by farmers to select sowing dates for cultivars that minimise frost or heat risk. (Zheng et al., 2013).

The research strategies outlined above were tried and tested in wheat and barley research over a period of more than two decades. Advances in genomics, phenomics and biological data-science allow an alternative "genome-to-phenome" approach to be applied to understand phenology and adaptation of oats (Figure 4). Core to this approach is to survey genetic diversity at high-resolution and genome-wide scale in a representative population that captures the flow of alleles through the history of oat breeding, spanning founders, international imports, key parents and modern elite cultivars. The advent of high-resolution oat genotyping-by-sequencing assays can facilitate this (Bekele et al., 2018; Bekele et al., 2020). Another approach is to generate population-scale transcriptomes from key organs or timepoints in the plant lifecycle. Transcriptome data can be used to detect SNP variation in the coding regions of genes that are likely to function in the relevant tissues/timepoints of interest and can also assay variation in gene expression states across the genome, in different accessions (Xu et al., 2015). By phenotyping the same panel under controlled conditions (long versus short days, with or without vernalization), together with industry relevant field trials, it is possible to link variation in phenology to chromosomal regions. This generates data at scale that can be used to resolve gene-by-environment interactions through the application of emerging analytical tools, such as machine learning. Access to marker-trait associations for phenology at

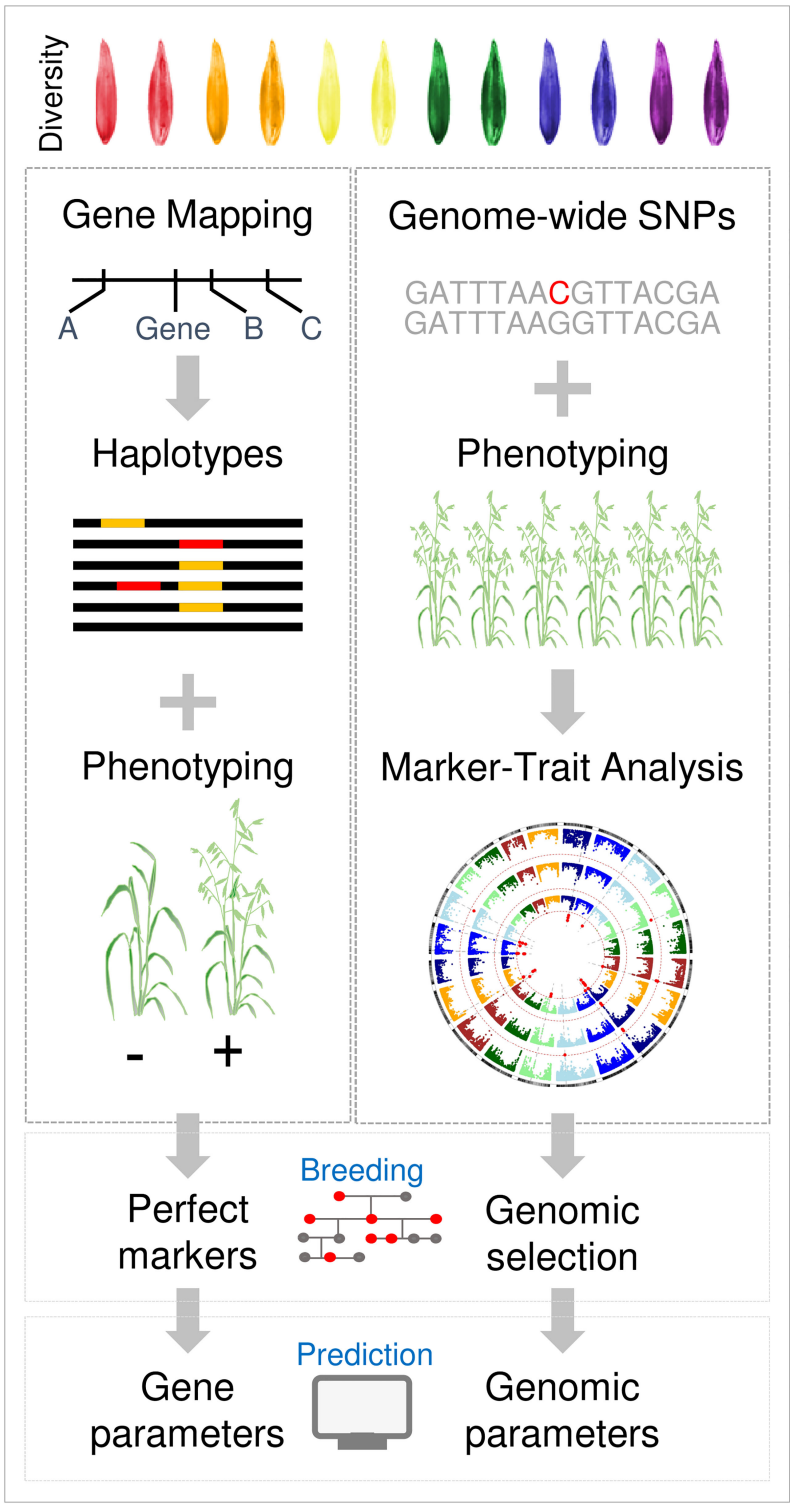


FIGURE 4
Overview of future oat phenology research strategies. A traditional strategy, based on approaches applied to wheat and barley, is compared with a "genome-to-phenome" strategy. In the traditional approach individual phenology traits are mapped to genes, using bi-parental mapping populations, for example. Knowledge of variation in these genes is then assembled by gene resequencing and used to design molecular markers for genotyping. In parallel, bespoke genetic resources such as near-isogenic lines or induced mutants are used to test gene function. In the genome-to-phenome strategy diversity panels are used to link genome-wide genetic diversity to broad phenotypic variation (phenology in this case). Marker-trait associations are then used for genomic prediction or to inform crop simulation models that predict phenology.

genome-wide scale in controlled environments that resolve variation for discrete aspects of phenology (i.e., vernalization and daylength requirements) also allows genomic parameterisation of simulation models or other data-driven prediction methods. Specifically, marker-trait association data can be used to set vernalization and daylength response parameters in APSIM to predict the flowering behaviour of individual oat cultivars. The same marker-trait association data can then be utilised in crop breeding through genomic selection/prediction. A "genome-to-phenome" strategy could be used alongside a more traditional approach and has the advantage that the genomic resources developed can be used to explore other aspects of oat biology.

Conclusions

Variation in phenology underlies adaptation of oat to climates, farming systems and end uses. Rapid progress in the development of genomic resourcess and the assembly of an oat reference genome sequence is accelerating progress towards understanding genes that underlie variation in phenology. Comparison of the physical locations of genetic markers linked to variation phenology with the likely locations of candidate phenology genes in the oat genome suggests that *VRN1* and *FT1* contribute to natural variation in oat phenology. There is now a need for detailed studies of the expression of these genes in diverse oat varieties, in response to different vernalization or photoperiod treatments, together with functional validation by mutagenesis or gene editing. Other future priorities will be to determine how haplotypes of phenology genes contribute to variation in oat phenology and to overall crop performance, by generating near-isogenic lines, for example. This knowledge can then be incorporated into genomic selection strategies, which can be used in crop breeding, and into phenology prediction tools that can be used to inform on-farm decision-making.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was jointly funded by the Grains Research and Development Corporation and CSIRO, project code CSP2007.

Acknowledgments

The authors wish to thank Dr Howard Eagles for his guidance and mentorship over several years. We also thank Dr Pamela Zwer and Dr Bruce Winter for providing pedigree information.

Conflict of interest

The authors declare that this study received funding from the Grains Research and Development Corporation (GRDC). The funder was not involved in the study design, collection, analysis interpretation of data, the writing of this article or the decision to submit it for publication.

Publisher's note

References

ABARES (2021). *Australian Crop report, Australian bureau of agricultural and resource economics and sciences*. (Canberra) doi: 10.25814/xqy3-sx57

Aitken, Y. (1974). *Flowering time, climate and genotype: the adaptation of agricultural species to climate through flowering responses* (Melbourne: Melbourne University Press).

Beales, J., Turner, A., Griffiths, S., Snape, J. W., and Laurie, D. A. (2007). A pseudo-response regulator is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (*Triticum aestivum* l.). *Theor. Appl. Genet.* 115, 721–733. doi: 10.1007/s00122-007-0603-4

Bekele, W. A., Itaya, A., Boyle, B., Yan, W., Fetch, J. M., and Tinker, N. A. (2020). A targeted genotyping-by-sequencing tool (RAPTURE) for genomics-assisted breeding in oat. *Theor. Appl. Genet.* 133, 653–664. doi: 10.1007/s00122-019-03496-w

Bekele, W. A., Wight, C. P., Chao, S., Howarth, C. J., and Tinker, N. A. (2018). Haplotype based genotyping-by-sequencing in oat genome research. *Plant Biotechnol. J.* 16, 1452–1463. doi: 10.1111/pbi.12888

Bell, G. H. D. (1936). Experiments on vernalization. *J. Agric. Sci.* 26:155–171. doi: 10.1017/S0021859600021869

Boden, S., Cavanagh, C., Cullis, B., Ramm, K., Greenwood, J., Finnegan, E. J., et al. (2015). *Ppd-1* is a key regulator of inflorescence architecture and paired spikelet development in wheat. *Nat. Plant* 1 (2), 14016. doi: 10.1038/nplants.2014.16

Boden, S. A., Weiss, D., Ross, J. J., Davies, N. W., Trevaskis, B., Chandler, P. M., et al. (2014). *EARLY FLOWERING3* regulates flowering in spring barley by mediating gibberellin production and *FLOWERING LOCUS t* expression. *Plant Cell* 26, 1557–1569. doi: 10.1105/tpc.114.123794

Bonnett, O. T. (1966). Development of the oat panicle, in inflorescences of maize, wheat, rye, barley and oats: their initiation and development. *Bulletin Illinois Agric. Exp. Station* 721, 92–102.

Borodin, D. N. (1934). *Yarovization formulas for winter oats* (New York). Published by the author.

Brassac, J., Muqaddsi, Q. H., Plieske, J., Ganal, M. W., and Röder, M. S. (2021). Linkage mapping identifies a non-synonymous mutation in *FLOWERING LOCUS T (FT-B1)* increasing spikelet number per spike. *Sci. Rep.* 11, 1585. doi: 10.1038/s41598-020-80473-0

Brown, P. D., McKenzie, R. I. H., and Mikaelsen, K. (1980). Agronomic, genetic, and cytologic evaluation of a vigorous new semidwarf oat. *Crop Sci.* 20, 303–306. doi: 10.2135/cropsci1980.0011183X002000030003x

Burrows, V. D. (1984). Donald Oats. *Can. J. Plant Sci.* 64, 411–413. doi: 10.4141/cjps84-059

Campoli, C., Pankin, A., Drosse, B., Casoa, C. M., Davis, S. J., and von Korff, M. (2013). *HvLUX1* is a candidate gene underlying the early maturity 10 locus in barley: phylogeny, diversity, and interactions with the circadian clock and photoperiodic pathways. *New Phytol.* 199, 1045–1059. doi: 10.1111/nph.12346

Chaffin, A. S., Huang, Y. F., Smith, S., Bekele, W. A., Babiker, E., Gnanesh, B. N., et al. (2016). A consensus map in cultivated hexaploid oat reveals conserved grass synteny with substantial subgenome rearrangement. *Plant Genome* 9 (2). doi: 10.3835/plantgenome2015.10.0102

Chen, Y., Carver, B. F., Wang, S., Zhang, F., and Yan, L. (2009). Genetic loci associated with stem elongation and winter dormancy release in wheat. *Theor. Appl. Genet.* 118, 881–889. doi: 10.1007/s00122-008-0946-5

Chen, A., and Dubcovsky, J. (2012). Wheat TILLING mutants show that the vernalization gene *VRN1* down-regulates the flowering repressor *VRN2* in leaves but is not essential for flowering. *PloS Genet.* 8, e1003134. doi: 10.1371/journal.pgen.1003134

Chen, A., Li, C., Hu, W., Lau, M. L., Lin, H., Rockwell, N. C., et al. (2014). PHYTOCHROME c plays a major role in the acceleration of wheat flowering under long day photoperiod. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10037–10044. doi: 10.1073/pnas.1409795111

Coblentz, W. K., Bertram, M. G., Martin, N. P., and Berzaghi, P. (2012). Planting date effects on the nutritive value of fall-grown oat cultivars. *Agron. J.* 104, 312–323. doi: 10.2134/agronj2011.0273

Collins, D., and King, G. D. (1798). "*An account of the English colony in new south wales. London,"* in: (London: Cadell and Davies in the Strand) Available at: http://gutenberg.net.au/ebooks/e00010.html.

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

Cuddeford, D. (1995). ""Oats for animal feed"," in *The oat crop: production and utilization*. Ed. R. W. Welch (Dordrecht: Springer), 321–368. doi: 10.1007/978-94-011-0015-1_11

Danyluk, J., Kane, N. A., Breton, G., Limin, A. E., Fowler, D. B., and Sarhan, F. (2003). TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. *Plant Physiol.* 132, 1849–1860. doi: 10.1104/pp.103.023523

Díaz, A., Zikhali, M., Turner, A. S., Isaac, P., and Laurie, D. A. (2012). Copy number variation affecting the *Photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in wheat (*Triticum aestivum*). *PloS One* 7 (3), e33234. doi: 10.1371/journal.pone.0033234

Dixon, L. E., Farré, A., Finnegan, E. J., Orford, S., Griffiths, S., and Boden, S. A. (2018). Developmental responses of bread wheat to changes in ambient temperature following deletion of a locus that includes *FLOWERING LOCUS T1*. *Plant Cell Environ.* 41, 1715–1725. doi: 10.1111/pce.13130

Dixon, L. E., Karsai, I., Kiss, T., Adamski, N. M., Liu, Z., Ding, Y., et al. (2019). *VERNALIZATION1* controls developmental responses of winter wheat under high ambient temperatures. *Development* 146, dev172684. doi: 10.1242/dev.172684

Dolferus, R., Ji, X., and Richards, R. A. (2011). Abiotic stress and control of grain number in cereals. *Plant Sci.* 181, 331–341. doi: 10.1016/j.plantsci.2011.05.015

Eagles, H. A., Cane, K., Kuchel, H., Hollamby, G. J., Vallance, N., Eastwood, R. F., et al. (2010). Photoperiod and vernalization gene effects in southern Australian wheat. *Crop Past. Sci.* 61, 721–730. doi: 10.1071/CP10121

Eagles, H. A., Cane, K., and Neil, V. (2009). The flow of alleles of important photoperiod and vernalisation genes through Australian wheat. *Crop Past. Sci.* 60, 646–657. doi: 10.1071/CP09014

Eagles, H. A., Cane, K., and Trevaskis, B. (2011). Veery wheats carry an allele of *Vrn-A1* that has implications for freezing tolerance in winter wheats. *Plant Breed.* 130, 413–418. doi: 10.1071/CP09014

Endo-Higashi, N., and Izawa, T. (2011). Flowering time genes *Heading date 1* and *Early heading date 1* together control panicle development in rice. *Plant Cell Physiol.* 52, 1083–1094. doi: 10.1093/pcp/pcr059

Esvelt Klos, K., Huang, Y. F., Bekele, W. A., Obert, D. E., Babiker, E., Beattie, A. D., et al. (2016). Population genomics related to adaptation in elite oat germplasm. *Plant Genome* 9 (2). doi: 10.3835/plantgenome2015.10.0103

Faure, S., Turner, A. S., Gruszka, D., Christodoulou, V., Davis, S. J., von Korff, M., et al. (2012). Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8328–8333. doi: 10.1073/pnas.1120496109

Feng, Y., Yin, Y., and Fei, S. (2017). *BdVRN1* expression confers flowering competency and is negatively correlated with freezing tolerance in *Brachypodium distachyon*. front. *Plant Sci.* 8. doi: 10.3389/fpls.2017.01107

Finkner, V. C., Ponoleit, C. G., and Davis, D. L. (1973). Heritability of rachis node number of *Avena sativa* l. *Crop Sci.* 13, 84–85. doi: 10.2135/cropsci1973.0011183X001300010026x

Finnegan, E. F., Ford, B., Wallace, X., Pettolino, F., Griffith, P., Schmitz, R. J., et al. (2018). Zebularine treatment is associated with deletion of *FT-B1* leading to an increase in spikelet number in bread wheat. *Plant Cell Environ.* 41, 1346–1360. doi: 10.1111/pce.13164

Fitzsimmons, R. W. (1988). Pridham, John Theodore (1879-1954) Australian Dictionary of Biography, Volume 11, 329 (Melbourne University Press).

Fitzsimmons, R. W., Roberts, G. L., and Wrigley, C. W. (1983). *Australian Oat varieties* (Melbourne: CSIRO Publishing). doi: 10.1071/9780643105447

Fjellheim, S., Boden, S., and Trevaskis, B. (2014). The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2014.00431

Flohr, B. M., Hunt, J. R., Kirkegaard, J. A., and Evans, J. R. (2017). Water and temperature stress define the optimal flowering period for wheat in south-eastern Australia. *Field Crops Res.* 209, 108–119. doi: 10.1016/j.fcr.2017.04.012

Fu, D. L., Szucs, P., Yan, L. L., Helguera, M., Skinner, J. S., von Zitzewitz, J., et al. (2005). Large Deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Mol. Gen. Genom.* 273, 54–65. doi: 10.1007/s00438-004-1095-4

Gasparis, S., Bregier, C., Orczyk, W., and Nadolska-Orczyk., A. (2008). Agrobacterium-mediated transformation of oat (*Avena sativa* l.) cultivars *via* immature embryo and leaf explants. *Plant Cell Rep.* 27, 1721–1729. doi: 10.1007/s00299-008-0593-y

Gauley, A., and Boden, S. A. (2021). Stepwise increases in *FT1* expression regulate seasonal progression of flowering in wheat (*Triticum aestivum*). *New Phytol.* 229, 1163–1176. doi: 10.1111/nph.16910

Gawroński, P., Ariyadasa, R., Himmelbach, A., Poursarebani, N., Kilian, B., Stein, N., et al. (2014). A distorted circadian clock causes early flowering and temperature-dependent variation in spike development in the eps-3Am mutant of einkorn wheat. *Genetics* 196, 1253–1261. doi: 10.1534/genetics.113.158444

Grau Nersting, L., Bode Andersen, S., von Bothmer, R., and Jorgensen, R. B. (2006). Morphological and molecular diversity of Nordic oat through one hundred years of breeding. *Euphytica* 150, 327–337. doi: 10.1007/s10681-006-9116-5

Harlan, J. R. (1982). ""Relationships between weeds and crops"," in *Biology and ecology of weeds*. Eds. M. Holzner and M. Numata (W. Junk Publishers: The Hague), 91–96.

Hemming, M. N., Fieg, S., Peacock, W. J., Dennis, E. S., and Trevaskis, B. (2009). Regions associated with repression of the barley (*Hordeum vulgare*) *VERNALIZATION1* gene are not required for cold induction. *Mol. Gen. Genom.* 282, 107–117. doi: 10.1007/s00438-009-0449-3

Hemming, M. N., Peacock, W. J., Dennis, E. S., and Trevaskis, B. (2008). Low-temperature and daylength cues are integrated to regulate *FLOWERING LOCUS t* in barley. *Plant Physiol.* 147, 355–366. doi: 10.1104/pp.108.116418

Herrmann, M. H., Yu, J., Beuch, S., and Weber, W. E. (2014). Quantitative trait loci for quality and agronomic traits in two advanced backcross populations in oat (*Avena sativa* l.). *Plant Breed.* 133, 588–601. doi: 10.1111/pbr.12188

Holland, J., Moser, H., O'Donoughue, L., and Lee, M. (1997). QTLs and epistasis associated with vernalization responses in oat. *Crop Sci.* 37, 1306–1316. doi: 10.2135/cropsci1997.0011183X003700040047x

Holland, J., Portyanko, A., Hoffman, L., and Lee, M. (2002). Genomic regions controlling vernalization and photoperiod responses in oat. *Theor. Appl. Genet.* 105, 113–126. doi: 10.1007/s00122-001-0845-5

Howarth, C. J., Martinez, P. M. J., Cowan, A. A., Griffiths, I. A., Lister, S. J., Langdon, T., et al. (2021). Genotype and environment affect the grain quality and yield of winter oats (*Avena sativa* l.). *Foods* 10, 2356. doi: 10.3390/foods10102356

Hunt, J., Lilley, J., Trevaskis, B., Flohr, B., Peake, A., Fletcher, A., et al. (2019). Early sowing systems can boost Australian wheat yields despite recent climate change. *Nat. Clim. Change* 9 (3), 244–247. doi: 10.1038/s41558-019-0417-9

Hyles, J., Bloomfield, M. T., Hunt, J. R., Trethowan, R. M., and Trevaskis, B. (2020). Phenology and related traits for wheat adaptation. *Heredity* 125, 417–430. doi: 10.1038/s41437-020-0320-1

Islam, M. R., Egrinya Eneji, A., Changzhomg, R., Yuegao, H., Gong, C., and Xuzhang, X. (2010). Oat-based cropping system for sustainable agricultural development in arid regions of northern China. *J. Agriculture Biotechnol. Ecol.* 3, 1–8.

Jenkins, G. (1973). The effect of sowing date and photoperiod on panicle morphology in naked oats (*Avena nuda*). *Ann. Appl. Biol.* 73, 85–94. doi: 10.1111/j.1744-7348.1973.tb01312.x

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., et al. (2022). The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature* 606, 113–119. doi: 10.1038/s41586-022-04732-y

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., et al. (1999). Activation tagging of the floral inducer *FT*. *Science* 286, 1962–1965. doi: 10.1126/science.286.5446.1962

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., et al. (2003). An overview of APSIM, a model designed for farming systems simulation. *Eur. J. Agron.* 18, 267–288. doi: 10.1016/S1161-0301(02)00108-9

King, S. R., and Bacon, R. K. (1992). Vernalization requirement of winter and spring oat genotypes. *Crop Sci.* 32, 677–680. doi: 10.2135/cropsci1992.0011183X003200030019x

Ki-Seung, K., Tinker, N., and Newell, M. (2014). Improvement of oat as a winter forage crop in the southern united states. *Crop Sci.* 54, 1336–1346. doi: 10.2135/cropsci2013.07.0505

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A pair of related genes with antagonistic roles in mediating flowering signals. *Science* 286, 1960–1962. doi: 10.1126/science.286.5446.1960

Lippi, M. M., Foggi, B., Aranguren, B., Ronchitelli, A., and Revedin, A. (2015). Multistep food plant processing at grotta paglicci (Southern Italy) around 32,600 cal B.P. *Proc. Natl. Acad. Sci. U.S.A.* 112, 12075–12080. doi: 10.1073/pnas.1505213112

Locatelli, A. B., Federizzi, L. C., Milach, S. C. K., Wight, C. P., Molnar, S. J., Chapados, J. T., et al. (2006). Loci affecting flowering time in oat under short-day conditions. *Genome* 49, 1528–1538. doi: 10.1139/g06-108

Loscos, J., Igartua, E., Contreras-Moreira, B., Gracia, M., and Casas, A. (2014). HvFT1 polymorphism and effect–survey of barley germplasm and expression analysis. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2

Matres, J. M., Hilscher, J., Datta, A., Armario-Nájera, V., Baysal, C., He, W., et al. (2021). Genome editing in cereal crops: an overview. *Transgenic Res.* 30, 461–498. doi: 10.1007/s11248-021-00259-6

Maughan, P. J., Lee, R., Walstead, R., Vickerstaff, R. J., Fogarty, M. C., Brouwer, C. R., et al. (2019). Genomic insights from the first chromosome-scale assemblies of oat (*Avena* spp.) diploid species. *BMC Biol.* 22, 92. doi: 10.1186/s12915-019-0712-y.014.00251

McKenzie, R. I. H., Brown, P. D., Martens, J. W., Harder, D. E., Nielsen, J., Gill, C. C., et al. (1984). Registration of Dumont oats. *Crop Sci.* 24, 207. doi: 10.2135/cropsci1984.0011183X002400010051x

McKeown, M., Schubert, M., Marcussen, T., Fjellheim, S., and Preston, J. C. (2016). Evidence for an early origin of vernalization responsiveness in temperate pooideae grasses. *Plant Physiol.* 172, 416–426. doi: 10.1104/pp.16.01023

Mengersen, F. (1960). Oat improvement in N.S.W. i. early history and development. II. development of new varieties in southern New South Wales. *Agric. Gazette NSW* 71, 449–461.

Mizuno, N., Nitta, M., Sato, K., and Nasuda, S. (2012). A wheat homologue of *PHYTOCLOCK 1* is a candidate gene conferring the early heading phenotype to einkorn wheat. *Genes Genet. Syst.* 87, 357–367. doi: 10.1266/ggs.87.357

Moore-Colyer, R. J. (1995). "Oats and oat production in history and pre-history"," in *The oat crop: production and utilization*. Ed. R. W. Welch (Dordrecht: Springer), 1–33. doi: 10.1007/978-94-011-0015-1_1

Murphy, J. P., and Hoffman, L. A. (1992). ""The origin, history, and production of oat"," in *Oat science and technology*. Eds. H. G. Marshall and M. E. Sorrells (Madison: Wiley), 1–28. doi: 10.2134/agronmonogr33.c1

Nava, I. C., Wight, C. P., Pacheco, M. T., Federizzi, L. C., and Tinker, N. A. (2012). Tagging and mapping candidate loci for vernalization and flower initiation in hexaploid oat. *Mol. Breed.* 30, 1295–1312. doi: 10.1007/s11032-012-9715-x

Nitcher, R., Pearce, S., Tranquilli, G., Zhang, X., and Dubcovsky, J. (2014). Effect of the hope FT-B1 allele on wheat heading time and yield components. *J. Heredity* 105, 666–675. doi: 10.1093/jhered/esu042

Oliver, S. N., Deng, W., Casao, M. C., and Trevaskis, B. (2013). Low temperatures induce rapid changes in chromatin state and transcript levels of the cereal *VERNALIZATION1* gene. *J. Exp. Bot.* 64, 2413–2422. doi: 10.1093/jxb/ert095

Peltonen-Sainto, P. (1994). Response to daylength in oats: pre-anthesis development and set of spikelets and florets. *J. Agron. Crop Sci.* 172, 104–112. doi: 10.1111/j.1439-037X.1994.tb0053x

Peltonen-Sainto, P., and Juahianen, L. (2020). Large Zonal and temporal shifts in crops and cultivars coincide with warmer growing seasons in Finland. *Regional Environ. Change* 20, 89. doi: 10.1007/s10113-020-01682-x

Peng, Y., Yan, H., Guo, L., Deng, C., Wang, C., Wang., Y., et al. (2022). Reference genome assemblies reveal the origin and evolution of allohexaploid oat. *Nat. Genet.* 54, 1248–1258. doi: 10.1038/s41588-022-01127-7

Portyanko, V. A., Chen, G., Rines, H. W., Phillips, R. L., Leonard, K. J., Ochocki, G. E., et al. (2005). Quantitative trait loci for partial resistance to crown rust, puccinia coronata, in cultivated oat, *Avena sativa* l. *Theor. Appl. Genet.* 111, 313–324. doi: 10.1007/s00122-005-2024-6

Preston, J. C., and Kellogg, E. A. (2008). Discrete developmental roles for temperate cereal grass *VERNALIZATION1/FRUITFULL*-like genes in flowering competency and the transition to flowering. *Plant Physiol.* 146, 265–276. doi: 10.1104/pp.107.109561

Pugsley, A. T. (1971). A genetic analysis of the spring-winter habit of growth in wheat. *Aust. J. Agric. Res.* 22, 21–31. doi: 10.1071/AR9710021

Ream, T. S., Woods, D. P., Schwartz, C. J., Sanabria, C. P., Mahoy, J. A., Walters, E. M., et al. (2014). Interaction of photoperiod and vernalization determines flowering time of *Brachypodium distachyon*. *Plant Physiol.* 164, 694–709. doi: 10.1104/pp.113.232678

Sampson, D. R., and Burrows, V. D. (1972). Influence of photoperiod, short-day vernalization, and cold vernalization on days to heading in *Avena* species and cultivars. *Can. J. Plant Sci.* 52, 471–482. doi: 10.4141/cjps72-077

Sasani, S., Hemming, M. N., Oliver, S., Greenup, A., Tavakkol-Afshari, R., Mahfoozi, S., et al. (2009). The influence of vernalization and daylength cues on the expression of flowering-time genes in the leaves and shoot apex of barley (*Hordeum vulgare*). *J. Exp. Bot.* 60, 2169–2178. doi: 10.1093/jxb/erp098

Shaw, P. D., Graham, M., Kennedy, J., Milne, I., and Marshall, DF (2014). Helium: visualization of large scale plant pedigrees. *BMC Bioinf.* 15, 259. doi: 10.1186/1471-2105-15-259

Sorrells, M. E., and Simmons, S. R. (1992). "Influence of environment on the development and adaptation of oat", in *Oat science and technology*. Eds. H. G. Marshall and M. E. Sorrells (Madison: Wiley), 115–163. doi: 10.2134/agronmonogr33.c1

Suarez-Lopez, 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

Sunstrum, F. G., Bekele, W. A., Wight, C. P., Yan, W., Chen, Y., and Tinker, N.A (2019). A genetic linkage map in southern-by-spring oat identifies multiple quantitative trait loci for adaptation and rust resistance. *Plant Breed.* 138, 82–94. doi: 10.1111/pbr.12666

Takahashi, Y., Shomura, A., Sasaki, T., and Yano, M. (2001). Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. *Proc. Natl. Acad. Sci. U.S.A.* 98, 7922–7927. doi: 10.1073/pnas.111136798

Tamaki, S., Matsuo, S., Wong, H. L., Yokoi, S., and Shimamoto, K. (2007). Hd3a protein is a mobile flowering signal in rice. *Science* 316, 1033–1036. doi: 10.1126/science.1141753

Tashiro, T., and Wardlaw, I. F. (1990). The response to high temperature shock and humidity changes prior to and during the early stages of grain development in wheat. *Aust. J. Plant Physiol.* 17, 551–561. doi: 10.1071/PP9900551

Tinker, N. A., and Deyl, J. K. (2005). A curated internet database of oat pedigrees. *Crop Sci.* 45, 2269–2272. doi: 10.2135/cropsci2004.0687

Tinker, N. A., Wight, C. P., Bekele, W. A., Yan, W., Jellen, E. N., Tsardakus Renhuldt, N., et al. (2022). Genome analysis in *Avena sativa* reveals hidden breeding barriers and opportunities for oat improvement. *Commun. Biol.* 5, 474. doi: 10.1038/s42003-022-03256-5

Trevaskis, B. (2018). Developmental pathways are blueprints for designing successful crops. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00745

Trevaskis, B., Bagnall, D. J., Ellis, M. H., Peacock, W. J., and Dennis, E. S. (2003). MADS box genes control vernalization-induced flowering in cereals. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13099–13104. doi: 10.1073/pnas.1635053100

Trevaskis, B., Hemming, M. N., Dennis, E. S., and Peacock, W. J. (2007). The molecular basis of vernalization-induced flowering in cereals. *Trends Plant Sci.* 12, 352–357. doi: 10.1016/j.tplants.2007.06.010

Turner, A., Beales, J., Faure, S., Dunford, R. P., and Laurie, D. A. (2005). The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* 310, 1031–1034. doi: 10.1126/science.1117619

Ubert, I. P., Zimmer, C. M., Pellizzaro, K., Federizzi, L. C., and Nava, I. C. (2017). Genetics and molecular mapping of the naked grains in hexaploid oat. *Euphytica* 213:41. doi: 10.1007/s10681-017-1836-1

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

von Zitzewitz, J., Szucs, P., Dubcovsky, J., Yan, L. L., Francia, E., Pecchioni, N., et al. (2005). Molecular and structural characterization of barley vernalization genes. *Plant Mol. Biol.* 59, 449–467. doi: 10.1007/s11103-005-0351-2

Wiggans, S. C., and Frey, K. J. (1955). Photoperiodism in oats. *Proc. Iowa Acad. Sci.* 62, 125–130. Available at: https://scholarworks.uni.edu/pias/vol62/iss1/13

Wight, C. P., Penner, G. A., O'Donoughue, L. S., Burrows, V. D., Molnar, S. J., and Fedak, G. (1994). The identification of random amplified polymorphic DNA

(RAPD) markers for daylength insensitivity in oat. *Genome* 37, 910–914. doi: 10.1139/g94-130

Woods, D. P., Ream, T. S., Bouché, F., Lee, J., Thrower., N., Wilkerson, C., et al. (2017). Establishment of a vernalization requirement in *Brachypodium distachyon* requires *REPRESSOR OF VERNALIZATION1*. *Proc. Natl. Acad. Sci. U.S.A.* 114, 6623–6628. doi: 10.1073/pnas.1700536114

Wooten, D. R., Livingston, D. P., Lyerly, H. J., Holland, J. B., Jellen, E. N., Marshall, D. S., et al. (2009). Quantitative trait loci and epistasis for oat winter-hardiness component traits. *Crop Sci*. 49, 1989–1998. doi: 10.2135/cropsci2008.10.0612

Xu, Q., Xing, S., Zhu, C., Liu, W., Fan, Y., Wang, Q., et al. (2015). Population transcriptomics reveals a potentially positive role of expression diversity in adaptation. *J. Int. Plant Biol.* 57, 284–299. doi: 10.1111/jipb.12287

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., et al. (2006). The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19581–19586. doi: 10.1073/pnas.0607142103

Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman., J., and Dubcovsky, J. (2004a). Allelic variation at the *VRN-1* promoter region in polyploid wheat. *Theor. Appl. Genet.* 109, 1677–1686. doi: 10.1007/s00122-004-1796-4

Yan, L., Loukoianov, A., Blechl, A., Tranquilli, G., Ramakrishna, W., SanMiguel, P., et al. (2004b). The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303, 1640–1644. doi: 10.1126/science.1094305

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6263–6268. doi: 10.1073/pnas.0937399100

Yan, W., Molnar, S. J., Fregeau-Reid, J., McElroy, A., and Tinker, N. A. (2007). Associations among oat traits and their responses to the environment. *J. Crop Improvement* 20, 1–29. doi: 10.1300/J411v20n01.01

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zakhrabekova, S., Gough, S. P., Braumann, I., Müller, A. H., Lundqvist, J., Ahmann, K., et al. (2012). Induced mutations in circadian clock regulator mat-a facilitated short-season adaptation and range extension in cultivated barley. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4326–4331. doi: 10.1073/pnas.1113009109

Zheng, B., Biddulph, B., Li, D., Kuchel, H., and Chapman, S. (2013). Quantification of the effects of *VRN1* and *Ppd-D1* to predict spring wheat (*Triticum aestivum*) heading time across diverse environments. *J. Exp. Bot.* 64, 3747–3761. doi: 10.1093/jxb/ert209

Zikhali, M., Wingen, L. U., and Griffiths, S. (2016). Delimitation of the *Earliness per se D1 (Eps-D1)* flowering gene to a subtelomeric chromosomal deletion in bread wheat (*Triticum aestivum*). *J. Exp. Bot.* 67, 287–929. doi: 10.1093/jxb/erv458

Zwer, P. (2017). Oats: grain-quality characteristics and management of quality requirements. *Cereal Grains* 10, 235–256. doi: 10.1016/B978-0-08-100719-8.00010-3

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 30 November 2022
DOI 10.3389/fpls.2022.987073

OPEN ACCESS

EDITED BY
Kyung Do Kim,
Myongji University, South Korea

REVIEWED BY
Hyun Jo,
Kyungpook National University, South Korea
Satoshi Watanabe,
Saga University, Japan

*CORRESPONDENCE
Fanjiang Kong
kongfj@gzhu.edu.cn
Baohui Liu
liubh@gzhu.edu.cn
Xiaoming Li
lionandmoon@scbg.ac.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 05 July 2022
ACCEPTED 24 October 2022
PUBLISHED 30 November 2022

CITATION
Lv T, Wang L, Zhang C, Liu S, Wang J, Lu S, Fang C, Kong L, Li Y, Li Y, Hou X, Liu B, Kong F and Li X (2022) Identification of two quantitative genes controlling soybean flowering using bulked-segregant analysis and genetic mapping.
Front. Plant Sci. 13:987073.
doi: 10.3389/fpls.2022.987073

Identification of two quantitative genes controlling soybean flowering using bulked-segregant analysis and genetic mapping

Tianxiao Lv[1†], Lingshuang Wang[1†], Chunyu Zhang[2†], Shu Liu[2,3†], Jinxing Wang[4], Sijia Lu[1], Chao Fang[1], Lingping Kong[1], Yunlong Li[4], Yuge Li[2], Xingliang Hou[2], Baohui Liu[1*], Fanjiang Kong[1*] and Xiaoming Li[2,3*]

[1]Guangdong Provincial Key Laboratory of Plant Adaptation and Molecular Design, Guangzhou Key Laboratory of Crop Gene Editing, Innovative Center of Molecular Genetics and Evolution, School of Life Sciences, Guangzhou Higher Education Mega Center, Guangzhou University, Guangzhou, China, [2]Key Laboratory of South China Agricultural Plant Molecular Analysis and Genetic Improvement, Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Innovative Academy of Seed Design, Chinese Academy of Sciences, Guangzhou, China, [3]University of Chinese Academy of Sciences, Beijing, China, [4]Suihua Branch Institute, Heilongjiang Academy of Agricultural Sciences, Suihua, Heilongjiang, China

Photoperiod responsiveness is important to soybean production potential and adaptation to local environments. Varieties from temperate regions generally mature early and exhibit extremely low yield when grown under inductive short-day (SD) conditions. The long-juvenile (LJ) trait is essentially a reduction and has been introduced into soybean cultivars to improve yield in tropical environments. In this study, we used next-generation sequencing (NGS)-based bulked segregant analysis (BSA) to simultaneously map qualitative genes controlling the LJ trait in soybean. We identified two genomic regions on scaffold_32 and chromosome 18 harboring loci *LJ32* and *LJ18*, respectively. Further, we identified *LJ32* on the 228.7-kb scaffold_32 as the soybean pseudo-response-regulator gene *Tof11* and *LJ18* on a 301-kb region of chromosome 18 as a novel *PROTEIN FLOWERING LOCUS T-RELATED* gene, *Glyma.18G298800*. Natural variants of both genes contribute to LJ trait regulation in tropical regions. The molecular identification and functional characterization of *Tof11* and *LJ18* will enhance understanding of the molecular mechanisms underlying the LJ trait and provide useful genetic resources for soybean molecular breeding in tropical regions.

KEYWORDS

long-juvenile, BSA, positional cloning, *Tof11*, *FT*

Introduction

Soybean [Glycine max (L.) Merr.], the main source of vegetable protein and oil globally, is a facultative short-day crop (GrahamVance, 2003). Flowering time and maturity traits significantly determine both plant adaptation to specific latitude and grain yield (Cober and Morrison, 2010; Zhong and Kong, 2022). Soybean is cultivated in a wide latitudinal range, from high-latitude areas such as Northeast China to tropical regions such as South America (Wang et al., 2016; Han et al., 2016). This broad ecological adaptability is enabled by genetic variation at major gene loci and quantitative trait loci (QTLs) controlling flowering and maturity (Lin et al., 2021a; Lin et al., 2021b). Multiple naturally occurring variants at these loci have become the targets of human selection and endow soybean with the flexibility to adapt to different areas with distinct photoperiod patterns. To date, 16 maturity loci, E1 to E11, J, Tof5, Tof11, Tof12, and LUX, have been identified by forward-genetic approaches (Bernard, 1971; Buzzell, 1971; McBlain and Bernard, 1987; Ray et al., 1995; Bonato and Vello, 1999; Cober and Voldeng, 2001a; Cober et al., 2010; Xia et al., 2012; Kong et al., 2014; Li et al., 2017; Samanfar et al., 2017; Wang et al., 2019; Lu et al., 2020; Bu et al., 2021; Dong et al., 2022). Among them, E1, E2, E3, E4, E7, E8, E10, Tof11, and Tof12 delay flowering and maturity under long-day (LD) conditions, and their recessive alleles enhance soybean adaptation to high latitudes (Cober and Voldeng, 2001b; Liu et al., 2008; Watanabe et al., 2009; Cober et al., 2010; Xia et al., 2012; Cao et al., 2017; Samanfar et al., 2017; Wang et al., 2019; Lu et al., 2020). A recent report indicated that the J protein associates with two LUX homologs to form the evening complex, which plays key roles in photoperiodic flowering and photoperiod sensitivity in soybean under both short-day (SD) and LD conditions (Bu et al., 2021). Tof11 and Tof12, encoding two homoeologous pseudo-response regulator (PRR) proteins, improved adaptation to the limited summer growth period at higher latitudes during soybean domestication (Lu et al., 2020).

At the other end of the latitudinal range, in the tropics, warm temperature and short photoperiod strongly induce rapid flowering and early maturity in photoperiod-sensitive soybean cultivars, making the vegetative phase very short and resulting in low yields (Parvez and Gardner, 1987; Destro et al., 2001). In these conditions, extension of the flowering and reproductive phases is necessary to allow greater vegetative growth and improve yield. The long-juvenile (LJ) trait has been introduced into tropical soybean cultivars to meet this need (Sinclair and Hinson, 1992; Carpentieri-Pípolo et al., 2002; Li et al., 2017; Lu et al., 2017). However, genetic information regarding this trait remains is limited. As the major classical locus conferring the LJ trait, J was identified as the ortholog of Arabidopsis thaliana EARLY FLOWERING 3 (ELF3). J depends genetically on the legume-specific flowering repressor E1 and directly downregulates E1 expression, thereby relieving the repression of two important FLOWERING LOCUS T (FT) genes (FT2a and FT5a) and promoting flowering under SD conditions (Lu et al., 2017; Fang et al., 2021). Recently, FT2a and FT5a were found to have variants of diverse origins that played distinctive roles as soybean spread to lower latitudes (Li et al., 2021). Tof16 was identified as a novel LJ locus that harbors the soybean homolog of the Arabidopsis LATE ELONGATED HYPOCOTYL (LHY), which delays flowering and improves yield at low latitudes (Dong et al., 2021). Loss of function of J, FT2a, or Tof16 is the major genetic base of soybean adaptation in tropical regions. Additionally, many QTLs associated with the LJ trait have been identified in soybean varieties (Fang et al., 2019; Lin et al., 2021a).

Conventional positional cloning and QTL mapping are powerful approaches for investigating the genetic control of phenotypic variation in agronomic traits (Burke et al., 2007). However, classical map-based gene cloning approaches are usually time-consuming owing to the need for genetic crossing and phenotypic analysis. As an alternative, the application of bulked segregant analysis (BSA) to QTL selection provides a simple strategy for rapidly identifying molecular markers tightly linked to the causal gene underlying a given phenotype (Giovannoni et al., 1991; Michelmore et al., 1991). BSA methods have been used in many organisms to map important genes (Mansur et al., 1993; Yi et al., 2006; Watanabe et al., 2011; Whipple et al., 2011). With the continuing advances in DNA sequencing technology, next-generation sequencing (NGS)-based BSA can dramatically accelerate the process of identifying causal genes of particular traits (Schneeberger and Weigel, 2011).

To investigate QTLs and corresponding candidate genes associated with the LJ trait, in this study we used genome-wide NGS-based BSA mapping of a soybean biparental population to identify two QTLs, named LJ32 and LJ18, conferring the LJ trait in soybean. We further validated these two QTLs and fine-mapped them by marker-based classical gene mapping to two intervals of 229 and 301 kb. Molecular and transgenic analyses demonstrated that the PRR gene Time of Flowering 11 (Tof11) and a PROTEIN FLOWERING LOCUS T-RELATED gene, Glyma.18G298800, may be responsible for the effects of the LJ32 and LJ18 loci. Overall, our study provides a useful genetic resource for soybean adaptation and molecular breeding to adapt soybeans for growth in tropical environments.

Results

Phenotypic analysis

To identify additional loci contributing to the LJ trait, we developed a set of 213 recombinant inbred line (RIL) populations from a cross between two closely related soybean cultivars. We used the near-isogenic lines (NILs) ZK193 and ZK158, both with the genetic background of the Canadian cultivar Harosoy from L62-812, which have the same genotypes (e1/e2/E3/E4/E9/Dt1) for the major flowering time

genes E1-E4 and E9 and the stem growth habit gene Dt1 (Supplementary Table 1). Nonetheless, ZK193 shows significantly earlier flowering and maturity than ZK158 under SD conditions (12 h light/12 h dark) (Figure 1). In addition, the two parents displayed different phenotypes in regard to several other traits, including node and pod number, grain number, and grain yield per plant (Supplementary Figure 1). To perform QTL mapping, we created cross combinations by pollinating ZK193 with pollen from ZK158 to develop descendant populations. We planted all RIL individuals and plants of both parent strains across 2 years (2018 and 2019) in Guangzhou, China, and recorded their flowering times at the R1 stage. The flowering data from 2018 and 2019 were strongly correlated ($P < 0.01$, $R = 0.724$), and we therefore used the data from 2019 for the following bulk segregation analysis (BSA) to detect flowering-associated loci.

Analysis of flowering-associated loci by BSA and QTL mapping

Based on the phenotypic assessment, we pooled genomic DNA from 30 individuals with extreme phenotypes (extremely early flowering and extremely late flowering) separately into an EF bulk sample and a LF bulk sample, respectively. We also extracted genome DNA from each parental line, isolated from leaves of 20 plants, for NGS-based BSA sequencing. After filtering, we identified clean reads and aligned them to the Williams 82 reference genome, and obtained high-quality single-nucleotide polymorphisms (SNPs) with which to calculate SNP-index values. We observed two peaks in the SNP-index plot, which we assigned as candidate flowering-time control regions in this population (Figure 2). The candidate regions were designated LJ trait 32 (LJ32) and LJ18 due to their locations on scaffold_32 and chromosome 18, respectively. SNP-index analysis revealed that the regions of the physical map around 229 kb on scaffold_32 and 4.7 Mb on chromosome 18 might be associated with flowering time (Supplementary Table 2, Supplementary Figure 2).

Scaffold_32 is a fragment that failed to be successfully assembled in the Williams 82 reference genome. Because it is a small fragment (229 kb), we performed ANOVA analysis to confirm the flowering-associated interval on Scaffold_32 (Supplementary Table 2). The results showed that the flowering times of all plants with theZK193 marker pattern for Scaffold_32 were significantly earlier than the mean value for plants with the ZK158 pattern. We inferred that a QTL is present in the same chromosome region as the marker Tof11, here named LJ32. Under SD conditions, LJ32 significantly affected flowering time, as shown by genome-wide analyses with permutation tests ($P < 0.05$) (Figure 2). To validate the candidate region chromosome 18 identified by BSA mapping, we used numerous insertion-deletion (Indel) and SNP markers in the region that are polymorphic between the two parental lines to genotype and constructed genetic linkage maps in the RIL population using the Kosambi function. A total of 11 markers, spanning 67.8 cM (Supplementary Figure 2), covered a part of the region of linkage group 18. The main marker type contributing to this linkage map was Indel markers, while the linkage gaps between the Indel markers were bridged by SNP markers. The constructed map was generally consistent with the US Department of Agriculture soybean genetic linkage map (Choi et al., 2007). QTL analyses revealed that LJ18 was located in a region between markers ID181191 and ID181120 on chromosome 18 (Supplementary Table 2). Two significant QTLs for flowering time, LJ32 and LJ18, were consistently detected in the 2018 and 2019 data, validating the accuracy of the NGS mapping results.

Characterization of LJ32

On the Scaffold_32 fragment, 22 genes were annotated. Of these, Glyma.U034500, encoding a PRR family protein, was

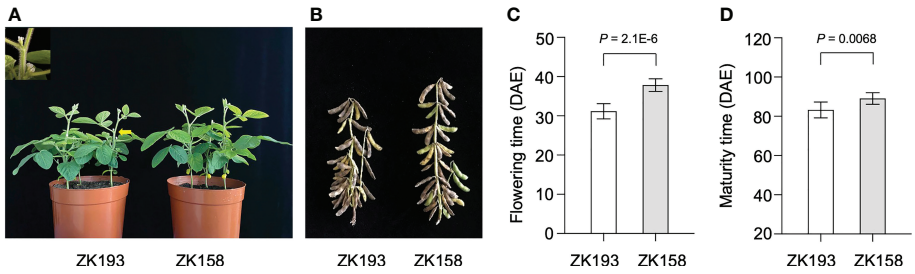


FIGURE 1
The phenotype of ZK193 and ZK158 under short-day (SD) conditions. **(A, B)** Flowering and maturity time of ZK193 and ZK158 under SD conditions. **(C)** Flowering time. Flowering time was recorded at the R1 stage. **(D)** Maturity time. Maturity time was recorded at the R8 stage. DAE (days after emergence). All data were given as mean ± SD (standard deviation, n = 10 plants). A Student's t-test was used to generate the P values.

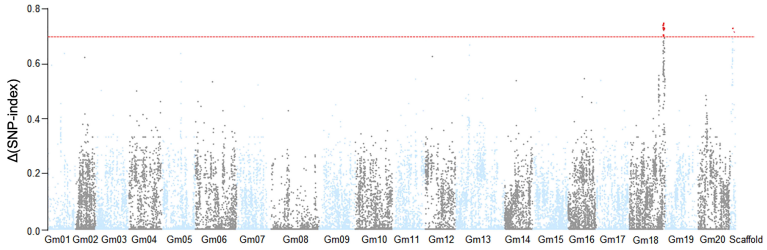


FIGURE 2
Identification of flowering time loci through SNP-index analysis. The result of SNP-index association analysis. The red lines show the threshold of Δ (SNP- index), which is represented by the top 5% of the permutation test. A larger value of Δ (SNP- index) indicates a stronger level of association. Under a threshold of 0.6964, two SNP markers on Gm18 and scaffold_32 significantly associated with the flowering time trait.

identified as Tof11, which has been reported to delay flowering in LD conditions (16 h light/8 h dark) (Lu et al., 2020) (Supplementary Table 3). Upon further analysis of the NGS data, we identified a 1-bp deletion (A2210-) in the last exon of Tof11, causing frameshifts and premature termination of protein translation, in the parent ZK193 compared with that in ZK158 (Supplementary Figure 3). Previous study found that this Tof11 haplotype, which we named Tof11-1, was the most abundant in landraces and improved cultivars and identified to be selected at an early stage of modern soybean breeding, and Tof11 is genetically dependent on E1 (Lu et al., 2020). We evaluated the effect of Tof11 on transcriptional regulation of E1 under SD conditions, A similar result was obtained in the parents showing that functional alleles of Tof11 in ZK158, relative to the respective mutant alleles in ZK193, increased E1 expression (Supplementary Figure 4). We thus identified Tof11 as a candidate gene potentially responsible for the effect of LJ32.

To characterize the function of Tof11 in soybean LJ regulation, we grew two complementary transgenic Tof11 lines (TC#2 and TC#4), along with a wild-type (WT) cultivar Dongnong 50 (Lu et al., 2020), under SD conditions. The TC#2 and TC#4 plants flowered slightly but significantly later than WT plants under SD conditions (Figure 3), supporting our hypothesis that LJ32 is encoded by TOF11.

To further validate the function of Tof11 in soybean LJ trait regulation under SD and identify the allelic variations of Tof11, we looked for variations in the Tof11 coding sequence in our collection of 338 re-sequenced soybean accessions from low-latitude regions grown in Guangzhou (Li et al., 2021). We identified 11 haplotypes, of which H2 and H4 (functional alleles) resulted in significantly later flowering than H1 and H3 (loss-of-function alleles) (Figure 4). The remaining alleles were not assessed because they were found in only a few accessions (Figure 4). Notably, in these 338 accessions, the frequency of functional alleles (48%) was similar to that of loss-of-function alleles, indicating that variations in Tof11 may contribute to the geographic distribution of soybean accessions in lower-latitude regions. Together, our observations indicated that Tof11 is the most likely causal gene in the LJ32 locus for the LJ trait.

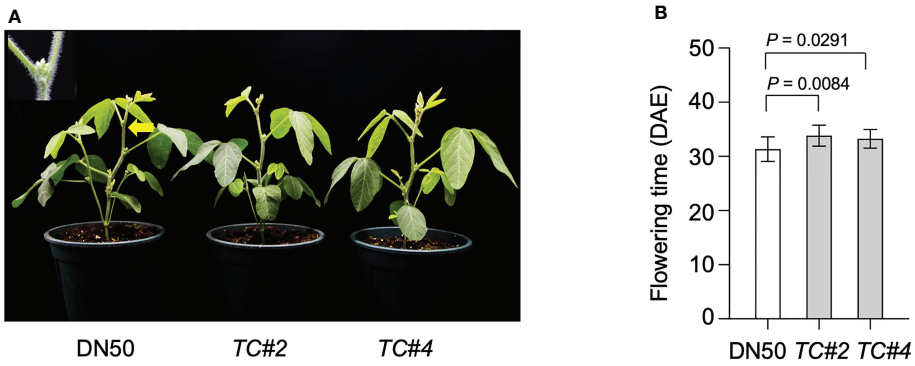


FIGURE 3
Confirmed identity of LJ32 by transgenic complementation. **(A)** The flowering phenotype of two independent transformants of complementation, TC#2 and TC#4, and the control, DN50 Tof11-1 under SD (12 h light/12 h dark). Scale bar, 10 cm. **(B)** Flowering time. All data were given as mean ± SD (standard deviation, n = 10 plants). A Student's t-test was used to generate the P values.

A

	Haplotype										Function
Zhonghuang 13	H1 (N=139)	C	G	T	G	G	GGCAA	G	G	-	-	+
	H2 (N=54)	C	G	T	G	G	GGCAA	G	G	A	-	-
	H3 (N=12)	G	G	T	G	G	GGCAA	G	G	A	-	-
	H4 (N=77)	C	G	T	G	G	GGCAA	G	C	A	-	-
	H5 (N=1)	C	G	A	G	G	GGCAA	G	G	A	-	-
	H6 (N=1)	C	A	T	G	G	GGCAA	G	C	A	-	-
	H7 (N=1)	C	G	T	T	G	GGCAA	G	C	A	-	-
	H8 (N=2)	C	G	T	G	A	GGCAA	G	C	A	-	-
	H9 (N=10)	C	G	T	G	A	-	G	C	A	-	-
	H10 (N=7)	C	G	T	G	G	GGCAA	A	C	A	-	-
	H11 (N=1)	C	G	T	G	A	GGCAA	G	C	A	A	-
CDS nt # in Zhonghuang 13		113	247	603	771	1063	1231	1384	1511	2210	2277	
AA # in Zhonghuang 13		38	79	201	257	355	409	462	504	737	759	
AA identified in Zhonghuang 13		S	V	L	V	V		A	G			
AA change		Stop	I	H	F	I	FS	T	A	FS	FS	

B

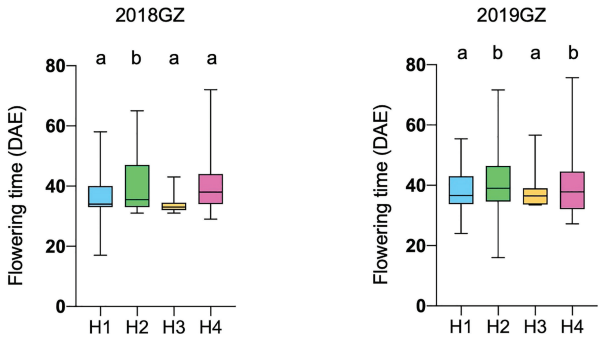


FIGURE 4
Flowering time of different alleles of Tof11. **(A)** Haplotypes and their origins of Tof11. FS, frameshift. **(B)** Flowering time of eleven haplotypes of Tof11 from the 338-accession panel. 20GZ, the accessions were planted in Guangzhou, China, in 2020; 18GZ, planted in Guangzhou, China, 2018. Two-tailed, two-sample t-test were used to generate the P values. Data represent mean ± SD of three biological replicates. Different lowercase letters represent significant differences at the level of $P < 0.05$, based on the Student's t test.

Positional cloning of LJ18

To further delineate the LJ18 locus, we surveyed the genotypes at two markers within the QTL in 1354 plants segregated from heterozygous plants and detected five recombinants. We also investigated the segregation pattern in the residual heterozygous lines (RHLs) (Supplementary Figure 5). Fine-mapping with seven additional molecular markers delimited the LJ18 genomic region to an ~301-kb region between markers M4 and M5 (Figure 5), which harbors 25 genes according to the Williams 82 reference genome (Supplementary Table 4). Among them, three PROTEIN FLOWERING LOCUS T-RELATED genes were annotated: Glyma.18G298800, Glyma.18G298900 (GmFT1a), and Glyma.18G299000 (GmFT1b) (Supplementary Figure 6). We found no variants in GmFT1a resulting in amino acid changes and one non-synonymous SNP each in GmFT1b and Glyma.18G298800 (LJ18) (Supplementary Figure 6, 7A).

GmFT1a and GmFT1b are widely recognized as FLOWERING LOCUS T (FT) homologs in soybean (Kong et al., 2014), and variants in promoter regions can regulate gene expression and function (Liu et al., 2018; Li et al., 2021). We therefore tested the functions of each gene in soybean flowering regulation using loss-of-function mutants generated by CRISPR/Cas9 gene editing in the Williams 82 cultivar. We obtained multiple mutants for both GmFT1a and GmFT1b (Supplemental Figures 7B, C). Unexpectedly, phenotypic analysis detected no significant difference in flowering time between the Williams 82 and these single mutants under SD conditions (Supplementary Figure 7D).

Another FT-related protein, encoded by Glyma.18G298800, was annotated and located in tandem with GmFT1a and GmFT1b on chromosome 18 (Supplementary Table 4). NGS data revealed a non-synonymous SNP in the 3rd exon specifying a polymorphism between the two parental strains at residue T170C (T57M), which is threonine (T) in ZK158 but methionine (M) acid in ZK193 (Figure 5C). These observations suggested that Glyma.18G298800 is a probable candidate for the causative gene in the LJ18 locus. We investigated the allelic variation of Glyma.18G298800 using the same strategy described above for Tof11 and identified five haplotypes, among which haplotype 2 (H2) corresponded to the lj18 allele (Figure 6A). we examined

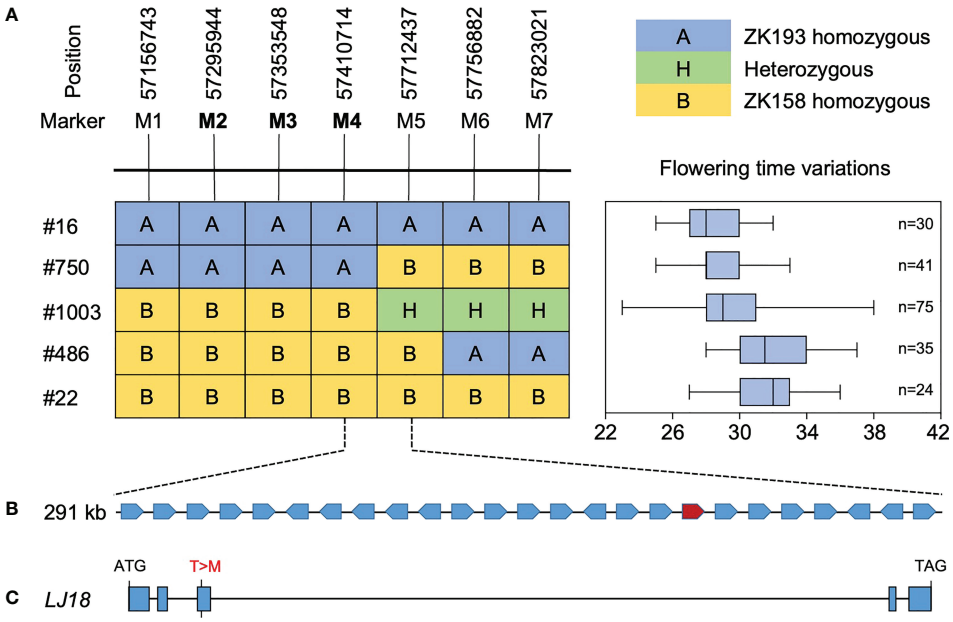


FIGURE 5
Identification of the LJ18 gene by map-based cloning. **(A)** Segregation of flowering time and delimitation of the LJ18 locus to a 301-kb region between the marker M4 and M5 on chromosome 18. A, homozygous for the allele from ZK193; B, homozygous for the allele from ZK158; H, heterozygous. Segregation of flowering time is shown in right box-plot format, where the interquartile region, median, and range are represented by the box, the bold vertical line, and the horizontal line, respectively. Each box plot corresponds to the segregant on the same row in left. **(B)** The delimited region contains 25 predicted genes. **(C)** Allelic variation in the LJ18 candidate gene Glyma.18G188800 between ZK193 and ZK158.

the geographic distribution of various alleles at LJ18 loci within the 338 accessions including the five LJ18 alleles (Figure 6A). All alleles showed no significant geographical distribution characteristics (Supplementary Table 5). We also analyzed the association of the Glyma.18G298800 haplotypes with flowering time in Guangzhou over 2 years. Accessions carrying H2 and H3 flowered significantly later than those carrying H1 (Figure 6B, C). This observation indicated that the polymorphism at nucleotide 170 in Glyma.18G298800 may lead to the variation in flowering time, supporting a role for Glyma.18G298800 in the control of flowering time in diverse genetic backgrounds. Taking these results together, we suggest that Glyma.18G298800 is a

A

	Haplotype				Function
Williams 82	H1 (n=177)	T	G	C	+
	H2 (n=80)	T	G	T	-
	H3 (n=33)	T	GTA	C	-
	H4 (n=1)	T	GTATA	C	-
	H5 (n=10)	C	GTATA	C	-
CDS nt # in Williams 82		26	121	170	
AA # in Williams 82		9	40	57	
AA identified in Williams 82		V		T	
AA change		A	Splicing	M	

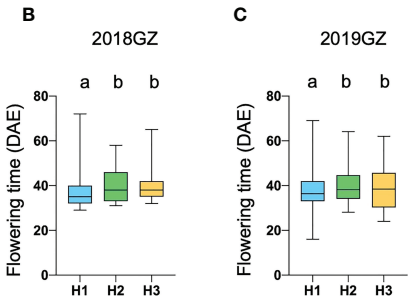


FIGURE 6
Flowering time variations of different alleles in LJ18. **(A)** Haplotypes and their origins of LJ18. **(B, C)** Variation of flowering time of LJ18 haplotypes from the 338-accession panel. 18GZ, the accessions were planted in Guangzhou, China, in 2018; 19GZ, planted in Guangzhou, China, 2019. Two-tailed, two-sample t-tests were used to generate the P values. Data represent mean ± SD of three biological replicates. Different lowercase letters represent significant differences at the level of $P < 0.05$, based on the Student's t test.

likely candidate gene for LJ18 and influences flowering regulation under SD conditions.

Functional analysis of Glyma.18G298800 in soybean flowering regulation

To confirm the expression patterns of Glyma.18G298800, we used reverse transcription quantitative PCR (RT-qPCR) to analyze the expression patterns of Glyma.18G298800 in leaves at different development stage of Williams 82 soybean (Supplemental Figure 8A). We found that Glyma.18G298800 transcripts were much more abundant in cotyledons and leaves than in other tissues. To preliminarily examine the function of Glyma.18G298800, we ectopically expressed Glyma.18G298800 in the Arabidopsis Columbia-0 (Col-0) ecotype. Among the resulting transgenic lines, RT-qPCR results showed that two independent lines exhibited relatively higher Glyma.18G298800 transcript levels than Col-0, and we selected these for further phenotypic analysis (Supplemental Figure 8B). The results showed that overexpression of Glyma.18G298800 significantly promoted flowering time compared with that of Col-0 (Figures 7A–C).

Next, we explored the molecular mechanism underlying the effect of Glyma.18G298800 on the LJ trait through a dual-luciferase assay. As soybean APETALA1 (AP1) is reported to be the primary target for FT regulation (Chen et al., 2020; Li et al., 2021), we used this assay to examine the effect of Glyma.18G298800 on AP1a, by fusing a 3-kb fragment of the AP1a promoter to the luciferase (LUC) reporter gene. We used Glyma.18G298800H1 and Glyma.18G298800H2 driven by the CaMV 35S promoter as effectors. We transformed each effector construct, together with the reporter construct, into N. benthamiana. Compared with the reporter vector control, the Glyma.18G298800H1 effector promoted the activity of AP1a promoter, as revealed by an increased LUC/REN (Figure 7D, E). The Glyma.18G298800H2 effector resulted in a lower LUC/REN ratio than Glyma.18G298800H1 effector, indicating that Glyma.18G298800H2 has no effect on the activity of the AP1a promoter (Figure 7D, E), which is consistent with the parental flowering phenotypes and with the results of the haplotype analysis above. Collectively, our results demonstrated that the PROTEIN

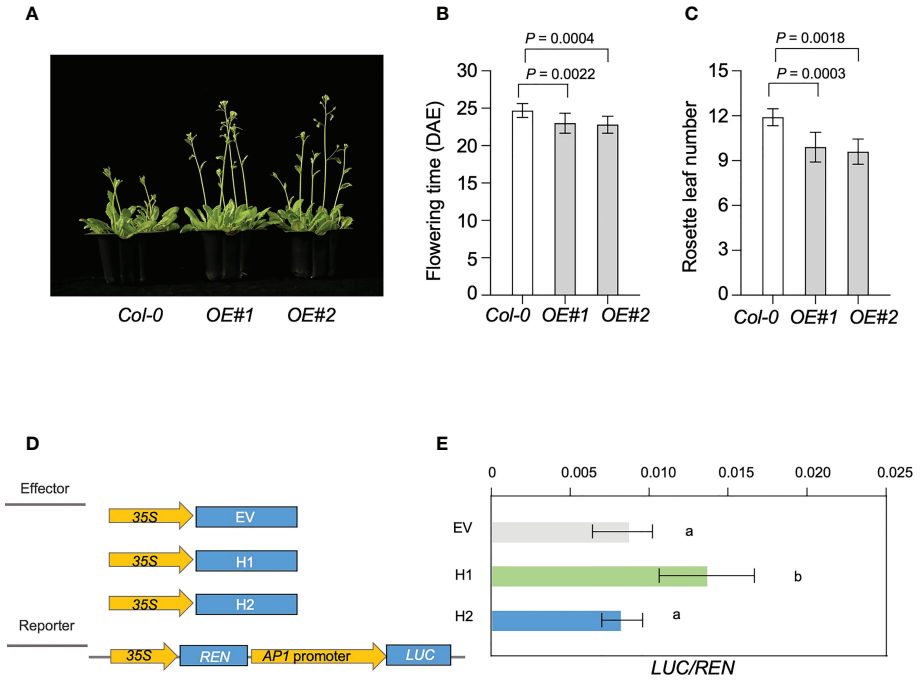


FIGURE 7
Genetic analysis of LJ18. **(A)** Flowering phenotype of the transgenic plants of Arabidopsis overexpressing LJ18, OE#1 and OE#2 relative to the untransformed control Col-0 plants. **(B)** Flowering time of Col-0, and LJ18-overexpressing plants under long day (LD) conditions at 22°C. **(C)** The number of rosette leaves in at least 10 plants at bolting was used as an indicator of flowering time. Statistically significant differences are indicated by different lowercase letters (Student's t test, $P < 0.05$). **(D)** Schematic maps of the constructs used in the transient expression assay. Empty vector (EV) was used as the negative control. Promoter of AP1a was separately inserted into the vector to get the reporters. 35S, CaMV35S promoter; LUC, firefly luciferase; REN, Renilla luciferase. **(E)** Transcriptional activity of AP1a promoter reporters in N. benthamina. REN was used as an internal control. LUC/REN represented the transcriptional activity of AP1a promoter. Data represent mean ± SD of three biological replicates. Different lowercase letters represent significant differences at the level of $P < 0.05$, based on the Student's t test.

FLOWERING LOCUS T-RELATED gene Glyma.18G298800 may function as a flowering promotor in soybean.

Discussion

In tropical regions, days are short during the growing season, neither temperature nor water are limiting, and the long-juvenile (LJ) trait is well established as an important adaptation that allows the crop to take full advantage of this favorable environment (Li et al., 2021). The introduction of the LJ trait in the 1970s overcame limitations on soybean growth, allowing its production to be extended to lower-latitude (tropical) areas (Neumaier and James, 1993; Carpentieri-Pípolo et al., 2002). For example, until 1960, soybean cultivars used in Brazil were imported from the United States, and cultivation areas were limited to above 22 degrees south latitude. In recent decades, however, the introduction of LJ germplasm has enabled Brazil to become the world's second-largest soybean producer (Neumaier and James, 1993). Notwithstanding the importance of LJ genes for soybean adaptation and yield in tropical regions, however, the underlying genetic basis and the trajectory of adaptation to low latitudes by means of these genes have remained largely unknown.

In this study, we developed a hybrid population from crosses between the NILs ZK193 and ZK158. NGS-based BSA combined with QTL analysis revealed two QTLs associated with the LJ trait, LJ32 and LJ18, located on scaffold_32 and chromosome 18, respectively, in the Williams 82 reference genome. Scaffold_32 is a 229-kb fragment that is assembled on chromosome 11 of the Zhonghuang 13 reference genome (Lu et al., 2020). We considered the gene Tof11, located in this region, as a candidate for the LJ32 locus. Although Tof11 is reported to delay flowering under LD conditions and to have contributed to ancient flowering-time adaptation (Lu et al., 2020), its function in LJ regulation had not previously been investigated. Here, we demonstrated that Tof11 functions as a flowering inhibitor under SD conditions and regulates the soybean LJ trait at low latitude.

FT proteins comprise a clade of the plant phosphatidylethanolamine-binding protein (PEBP) family and act as highly conserved regulators that are pivotal in the flowering pathways of various crop species (Kong et al., 2010). In soybean, several FT homologs have been reported as candidate florigens or antiflorigens under LD or SD conditions, including GmFT2a, GmFT5a, GmFT1a, and GmFT4 (Zhao et al., 2016; Liu et al., 2018; Jiang et al., 2019; Lu et al., 2020). Furthermore, GmFT2a and GmFT5a act as floral promoters conferring the LJ trait and played distinct roles as soybean spread to lower latitudes (Nan et al., 2014; Cai et al., 2020; Li et al., 2021; Yue et al., 2021). Here, we report that Glyma.18G298800, a novel FT homolog, is a possible candidate for the LJ18 locus that regulates the soybean LJ trait. This finding provides preliminary evidence that Glyma.18G298800 may contribute to delaying the flowering time of soybean varieties by inhibiting AP1a expression.

In summary, we used NGS-based BSA combined with QTL analysis to reveal two different QTLs conferring the LJ trait. We identified LJ32 as the soybean PRR gene Tof11 and LJ18 as the PROTEIN FLOWERING LOCUS T-RELATED gene Glyma.18G298800. The natural variants of both genes have significant influence on flowering time in SD accessions, suggesting that these two genes may play important roles in controlling flowering time in tropical regions. The identification and characterization of these LJ-related genes will contribute to the understanding of the genetic and molecular mechanisms underlying the LJ trait and could be used to ensure the successful deployment of high-yield germplasm in tropical environments.

Materials and methods

Plant materials, growth conditions, and phenotyping

The NILs lines, ZK193 and ZK158 with genetic background of Canadian cultivar Harosoy from L62-812 were used; they have the same maturity genotypes at E1, E2, E3, and E4 (e1/e1 e2/e2 E3/E3 E4/E4 DT1/DT1) (Supplementary Table S1). The cross combinations were made by pollinating ZK193 with pollen from ZK158 to develop RIL populations that were used for (NGS)-based Bulked segregation analysis (BSA) (Supplementary Table S2). The populations and low-latitude-adapted accessions for phenotyping were grown under naturally SD conditions (12 h light/12 h dark) in the field from 2017 to 2018 at the experimental station of Guangzhou University, Guangzhou (22° 26′ N, 112° 57′ E). The transformants for phenotyping were sown in pots in growth cabinets under SD conditions (12 h light/12 h dark)

Days to flowering were recorded at the R1 stage (first open flower appeared) for each plant47. The R1 values reported for the parents and populations are means from 10 plants. The number of parents and RIL plants used in each experiment are listed in Supplementary Table 2. Days to flowering (R1) were individually recorded and subjected to analysis of variance. Means of days to flowering among lines were compared with Tukey's HSD test using the Statistica software 03J (StatSoft). Plant height, number of branches, number of nodes, number of pods per plant, number of grains per plant, and yield per plant were all recorded at the R8 stage. All data are given as means ± s.e.m. (n = 10 plants). Two-tailed, two-sample t-tests were used to generate the P values.

DNA bulks construction and illumina sequencing

BSA was used to group the RIL population and its parents, Two DNA bulks for sequencing were first made by selecting

extreme individuals from the 213 RIL population plants with the basic statistics of the phenotypic data. One pool for early flowering comprised 30 lines with early flowering time and the other pool for late flowering involved 21 lines. DNA was extracted individually from leaves of plants, using a genomic DNA purification kit (Thermo Fisher Scientific Inc., United States) according to the manufacturer's protocol. The GC content, repeated sequences, and genetic characteristics of the DNA pools were analyzed by Biomarker (Beijing, China). About 20 μg of DNA from the two bulks and two parental lines were used to construct paired-end sequencing libraries, The genomic DNA pools were digested using the XhoI and MseI restriction enzymes (NEB, Ipswich, MA, USA), followed by PCR amplification, fragment amplification, fragment selection, fragment extraction and amplification, and fragment sequencing using the Illumina HiseqTM 2500 (Illumina, Inc; San Diego, CA, USA) at Biomarker Technologies. Real-time monitoring was performed for each cycle during sequencing and the ratio of high-quality reads with quality scores greater than Q30 (a quality score of 30, indicating a chance of 0.1% for an error and thus 99.9% confidence) in the raw reads GC content was calculated for quality control.

After removing adapter and low-quality reads, the clean reads were further rechecked for quality control using FASTQC1. High-quality sequences were aligned and to the Glycine max Wm82.a2.v1 reference genome from Phytozome2 (https://phytozome-next.jgi.doe.gov) using BWA with default parameters (Langmead and Salzberg, 2012).

SNP-index association analysis

GATK (Genome Analysis Toolkit) was used to call SNPs and small indels across parental lines and bulks (McKenna et al., 2010). The relative marker abundance in bulked DNA pool 1 (the early flowering pool) was calculated as the number of reads of the paternal allele divided by the total of reads which then gives proportion paternal alleles (or alternatively maternal alleles), whereas in pool 2 (the late flowering pool), Homozygous SNPs between parental lines and high-quality SNPs (minimum sequence read depth: 10 with SNP base quality ≥ 100 in pools) were selected for SNP-index analysis. A SNP-index was calculated at each SNP position for both pools using the base in parental lines as alternative base (Abe et al., 2012; Takagi et al., 2013b). Thus, the SNP-index was assigned as 0 or 1, when entire short sequence reads contained genomic fragments derived from parental lines, respectively. A Δ(SNP-index) was calculated by subtraction of the early flowering index from the late flowering index (Fekih et al., 2013; Takagi et al., 2013a). Thus, a high Δ(SNP-index) value of a SNP locus is indicative of an allele that was both very frequent in the pool 1and depleted in the pool 2.

QTL identification and statistical analysis

The polymorphisms between the parents introduced two kinds of markers, SNP and Indel. Markers were developed based on re-sequencing data from the parents, ZK193 and ZK158. The whole genome re-sequencing of ZK193, ZK158 and the Indel analysis using the software of SOAPindel was conducted by BGI-Shenzhen, China as described previously (Kong et al., 2014). The procedures for polymerase chain reaction and gel-electrophoresis were adopted as reported earlier (Li et al., 2017). Marker order and distance were determined by Map Manager Program QTXb20 using the Kosambi function and a criterion of 0.001 probability (d.f. = 1) and a genetic map was constructed (Lu et al., 2015). Mapchart 2.1 was used to draw the linkage groups (Voorrips, 2002). The Multiple QTL Model (MQM), implemented by MapQTL 5.0 was used for QTL detection (Van Ooijen, 2004). A LOD score of 2.5 was used as a minimum to declare the significance of a QTL in a particular genomic region. The tests of 1000 permutations at a 0.05 probability were conducted to identify the genome-wide LOD score (Churchill and Doerge, 1994).

Resequencing and variation calling

The resequencing data, VCF files and flowering time data from the 338-accession panel used in this study were obtained from Li et al. (Li et al., 2021). The VCF files were processed using the VCFtools software (v.0.1.16). Paired-end resequencing reads of the 338 accessions were mapped to the Glycine max Wm82.a2.v1 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Gmax) with BWA software (Version: 0.7.17-r1188, http://bio-bwa.sourceforge.net/) with the default parameters. The duplicates of sequencing read for each accession were filtered with the Picard package (Version: V1.109, http://broadinstitute.github.io/picard/), and uniquely mapping reads were retained in BAM format. Reads around indels from the BWA alignment were realigned with the IndelRealigner option in the Genome Analysis Toolkit (GATK, Version: V3.2-2, https://gatk.broadinstitute.org/hc/en-us). SNP and indel calling were performed with GATK and SAMtools software (Version: 1.9, http://samtools.sourceforge.net/). SNPs with MAFs less than 1% were discarded, and indels with a maximum length of 20 bp were included. SNP annotation was carried out based on the Williams82 genome, with Annovar (https://annovar.openbioinformatics.org/en/latest/).

Fine Mapping of LJ18 locus

For genetic analysis, we surveyed the genotypes by two markers for the F4 populations and conducted a segregating-heterozygous inbred family (n = 1354) that was heterozygous at LJ18 locus. The segregation pattern was carefully observed in the residual heterozygous lines (RHLs), in which the segregation

occurred only at LJ18 locus, seven additional Indel and SNP markers between markers ID181191 and ID181120 were identified (Figure 5A; Supplementary Table S5). Three recombinants in the region between M1 and M7 were genotyped using four Indel markers and three SNP markers (bold) (Figure 5A), and the flowering time of their progenies were evaluated to delimit the genomic interval containing LJ18. The genotypes of the LJ18 allele were analyzed by tagging marker M4 or M6 (Supplementary Table S5). The LJ18 allele was genotyped by its functional markers.

Plasmid construction and plant transformation

The CDS of the LJ18 candidate gene Glyma.18G298800 were obtained from ZK193 and ZK158. The CDS fragments were amplified by overlapping PCR to obtain one fragment and then introduced into the PTF101-Gene vector (containing the bar gene for glufosinate resistance) (Li et al., 2021). The construct (PTF101-35S:LJ18) was next introduced into the Agrobacterium tumefaciens strain EHA105, and 35S:LJ18-3flag transgenic lines were obtained through A. tumefaciens mediated transformation using the floral dip method in Col-0 wild-type followed by screening with 1/500 10% (w/v) basta. The FT1b knockout construct was produced by CRISPR-Cas9 as described previously (Ma et al., 2015). Two 20-bp sequences in the exons of FT1a and FT1b were selected as target sites for Cas9 cleavage (Supplementary Figure S6). Primers used for plasmid construction are listed in Supplementary Table S6. The above-mentioned CRISPR-Cas9 plasmid was transformed into Williams82 plants, and the transgenic plants were selected by basta (Ingbio, Lot: CB26213210).

Gene expression analysis

Soybean seedlings grown under SD conditions were harvested from the leaf of V3-stages for total RNA extraction using E.Z.N.A. Total RNA Kit I (Omega) and reverse transcribed to cDNA using MMLV-Reverse Transcriptase (Promega). qPCR was performed using a LightCycler 480 thermal cycler system (Roche) with KAPA SYBR Fast qPCR Kit Master Mix (Kapa Bio). The difference between the cycle threshold of target genes and the cycle threshold of the control gene was calculated by the relative quantification method ($2^{-\triangle\triangle}$Ct) and used to evaluate quantitative variation (Livak and Schmittgen, 2001). All expression analyses were performed with at least three biological replicates (three replicates of samples were taken from the same batch of plants, and total RNA was extracted from the pooled three to five tissues per independent replicate). The above experiments were independently performed at least three times, and representative results are shown. The primers used for gene expression analysis are listed in Supplementary Table S6.

Transient expression assays

To generate the AP1a pro-LUC reporter construct, ~3 kb AP1a promoter was cloned into the pGreenII0800-LUC vector (Li et al., 2021). The Renilla Luciferase (REN) gene under the control of 35S promoter in the pGreenII0800-LUC vector was used as the internal control. The coding regions of 35Spro: LJ18-H1 and 35Spro : LJ18-H2 were cloned into the pGreenII62-SK vector and used as effectors. All primers used for these constructs are listed in Supplementary Table S6. The A. tumefaciens mixtures were infiltrated into three leaves of tobacco plants as described previously (Yue et al., 2021). p35S-LJ18-H1-LUC and p35S-LJ18-H2-LUC represent the A. tumefaciens carrying the effector constructs and the control vector pGreenII 0800-LUC. The tobacco leaves were allowed to recover for 48 h. LUC/REN activity showing the results from three independent replications and the value of each replication were represented by a dot. The LUC and REN activities were measured using the Dual-Luciferase Reporter Assay System (Promega) under the manufacturers' instructions. The LUC/REN ratio was presented with three biological replicates.

Data availability

For phenotypic evaluation, at least ten individual plants were analyzed per accession, and the exact number of individuals (n) are presented in all figure legends. The exact number of replicates is given in figure legends. Mean values for each measured parameter were compared using one-way analysis of variance from SPSS (version 20, IBM) or one-tailed, two-sample Student's t-tests from Microsoft Excel, whenever appropriate; the statistical tests used for each experiment are given in the figure legends. Whole-genome sequencing data for ZK193, ZK158, and the two bulks are deposited at CNCB-NGDC and are publicly available as of the date of publication. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Data availability statement

The data presented in the study are deposited in the "Bioproject" repository accession number "PRJNA896173".

Author contributions

XL, FK, and BL designed the experiments, supervised the study, and managed the projects. TL, LW, CZ, and ShL. performed most of the research. JW, SiL, CF, LK, YunL, YugL, and XH performed data analysis. TL, CZ, SL, and XL drafted and revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the National Natural Science Foundation of China (31801390, 32090064, 31725021, 32201800). This work was also supported by Major Program of Guangdong Basic and Applied Research (2021A1515110522), Guangdong Provincial Key Laboratory of Plant Adaptation and Molecular Design (2022B1212010013-3) and Hainan Yazhou Bay Seed Lab of China (B21HJ0110).

Acknowledgments

We would like to acknowledge Dr. Shulin Liu for the data analysis.

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.2022.987073/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Yield related traits of the parent lines. **(A–F)** Plant height, number of nodes, number of branches, total pods per plant, total seeds per plant, and 100 seeds weight of ZK193 and ZK158. The plants were grown in standard field under artificially controlled SD (12 h light/12 h dark). All data are given as mean ± s.e.m. (n = 10 plants). Two-tailed, two-sample *t*-tests were used to generate the *P* values.

SUPPLEMENTARY FIGURE 2
Linkage groups containing QTLs for flowering time found in RIL populations. The Linkage groups containing QTLs for flowering time found in RIL populations. The title of linkage groups followed by the chromosome number in parenthesis is indicated at the top. The genetic distance of markers (cM) from the top of each linkage group are given on the left-hand side. The arrows indicate the position of the QTL peaks.

SUPPLEMENTARY FIGURE 3
Protein sequence alignment of *Tof11* alleles. Similarity of the predicted amino acid sequence of Tof11 homologs. The sequences were extracted from the Zhonghuang 13 genome database.

SUPPLEMENTARY FIGURE 4
The expression level of *E1* in ZK193 and ZK158. Expression level of *E1* in the NILs. Soybean *Tublin* was used as an internal control. Data are means ± s.e.m (n = 3). A Student's *t*-test was used to generate the *P* values.

SUPPLEMENTARY FIGURE 5
Frequency distribution of flowering times in *LJ18* residual heterozygous lines (RHL) populations.

SUPPLEMENTARY FIGURE 6
Protein sequence alignment of FT homologs. Similarity of the predicted amino acid sequences of FT homologs in soybean. The sequence of *LJ18* (Glyma.18G298800), *FT1a* (Glyma.18G298900), *FT1b* (Glyma.18G299000), *FT2a* (Glyma.16G150700), *FT2b* (Glyma.16G151000), *FT3a* (Glyma.16G044200), *FT3b* (Glyma.19G108100), *FT4* (Glyma.08G363100), *FT5a* (Glyma.16G044100), *FT5b* (Glyma.19G108200), and *FT6* (Glyma.08G363200) were obtained from the Williams 82 genome database. Highly conserved amino acids are in dark blue, blue, and light blue depending on the level of identity (darker = higher level). The light-yellow arrows indicate the Phosphatidyl Ethanolamine-Binding Protein (PEBP) domain (predicted by NCBI CD-Search: https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The red box represents the site of core amino acid replacement in the PEBP domain of lj18.1_H2.

SUPPLEMENTARY FIGURE 7
Plant material for experiments. **(A)** The predicted amino acid sequence of Glyma.18G299000 (*FT1b*) in ZK193 and ZK158, obtained from the Williams 82 genome database**. (B, C)** Characterization of *ft1a* and *ft1b* mutants without *Cas9* gene in the T_2 generation. Sequences of WT and mutant plants at target sites. Dashes indicate deleted nucleotides. Nucleotides in red indicate PAM. Red arrowheads indicate mutation locations. **(D)** Flowering time of Cas9-FT1a and Cas9-FT1b under SD (12 h light/12 h dark) conditions.

SUPPLEMENTARY FIGURE 8
Expression pattern of *Glyma.18G298800*. **(A)** Expression of *Glyma.18G298800* in difference tissues of Williams82 under SD. Soybean *Tublin* was used as an internal control. Data are means ± s.e.m (n = 3). **(B)** Expression level of *Glyma.18G298800* in the transgenic lines of Arabidopsis. OE, overexpressing. *Arabidopsis UBQ10* was used as an internal control. Data are means ± s.e.m (n = 3).

References

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., et al. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* 30, 174–178. doi: 10.1038/nbt.2095

Bernard, R. (1971). Two major genes for time of flowering and maturity in soybeans. *Crop Sci.* 11, 242–244. doi: 10.2135/cropsci1971.0011183X001100020022x

Bonato, E. R., and Vello, N. A. (1999). *E6*, a dominant gene conditioning early flowering and maturity in soybeans. *Genet. Mol. Biol.* 22, 229–232. doi: 10.1590/S1415-47571999000200016

Bu, T. T., Lu, S. J., Wang, K., Dong, L. D., Li, S. L., Xie, Q. G., et al. (2021). A critical role of the soybean evening complex in the control of photoperiod sensitivity and adaptation. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2010241118. doi: 10.1073/pnas.2010241118

Burke, J. M., Burger, J. C., and Chapman, M. A. (2007). Crop evolution: from genetics to genomics. *Curr. Opin. Genet. Dev.* 17, 525–532. doi: 10.1016/j.gde.2007.09.003

Buzzell, R. I. (1971). Inheritance of a soybean flowering response to fluorescent-daylength conditions. *Can. J. Genet. Cytol.* 13, 703–707. doi: 10.1139/G71-100

Cai, Y. P., Wang, L. W., Chen, L., Wu, T. T., Liu, L. P., Sun, S., et al. (2020). Mutagenesis of GmFT2a and GmFT5a mediated by CRISPR/Cas9contributes for expanding the regional adaptability of soybean. *Plant Biotechnol. J.* 18, 298–309. doi: 10.1111/pbi.13199

Cao, D., Takeshima, R., Zhao, C., Liu, B. H., Jun, A., and Kong, F. J. (2017). Molecular bases of flowering under long days and stem growth habit in soybean. *J. Exp. Bot.* 68, 1873–1884. doi: 10.1093/jxb/erw394

Carpentieri-Pípolo, V., Alves de Almeida, L., and de Souza Kiihl, R. A. (2002). Inheritance of a long juvenile period under short-day conditions in soybean. *Genet. Mol. Biol.* 25, 463–469. doi: 10.1590/S1415-47572002000400016

Chen, L. Y., Nan, H. Y., Kong, L. P., Yue, L., Yang, H., Zhao, Q. S., et al. (2020). Soybean AP1 homologs control flowering time and plant height. *J. Integr. Plant Biol.* 62, 1868–1879. doi: 10.1111/jipb.12988

Choi, I. Y., Hyten, D. L., Matukumalli, L. K., Song, Q. J., Chaky, J. M., Quigley, C. V., et al. (2007). A soybean transcript map: gene distribution, haplotype and single-nucleotide polymorphism analysis. *Genetics.* 176, 685–696. doi: 10.1534/genetics.107.070821

Churchill, G. A., and Doerge, R. (1994). Empirical threshold values for quantitative trait mapping. *Genetics.* 138, 963–971. doi: 10.1093/genetics/138.3.963

Cober, E. R., Molnar, S. J., Charette, M., and Voldeng, H. D. (2010). A new locus for early maturity in soybean. *Crop Sci.* 50, 524–527. doi: 10.2135/cropsci2009.04.0174

Cober, E. R., and Morrison, M. J. (2010). Regulation of seed yield and agronomic characters by photoperiod sensitivity and growth habit genes in soybean. *Theor. Appl. Genet.* 120, 1005–1012. doi: 10.1007/s00122-009-1228-6

Cober, E. R., and Voldeng, H. D. (2001a). A new soybean maturity and photoperiod-sensitivity locus linked to *E1*. *Crop Sci.* 41, 698–701. doi: 10.2135/cropsci2001.413698x

Cober, E. R., and Voldeng, H. D. (2001b). Low R:FR light quality delays flowering of E7E7 soybean lines. *Crop Sci.* 41, 1823–1826. doi: 10.2135/cropsci2001.1823

Destro, D., Carpentieri-Pipopo, V., de Souza Kiihl, R. A., and Alves de Almeida, L. (2001). Photoperiodism and genetic control of the long juvenile period in soybean: a review. *Crop Breed. Appl. Biotech.* 1, 72–92. doi: 10.13082/1984-7033.v01n01a10

Dong, L. D., Cheng, Q., Fang, C., Kong, L. P., Yang, H., Hou, Z. H., et al. (2022). Parallel selection of distinct Tof5 alleles drove the adaptation of cultivated and wild soybean to high latitudes. *Mol. Plant* 15, 308–321. doi: 10.1016/j.molp.2021.10.004

Dong, L. D., Fang, C., Cheng, Q., Su, T., Kou, K., Kong, L. P., et al. (2021). Genetic basis and adaptation trajectory of soybean from its temperate origin to tropics. *Nat. Commun.* 12, 5445. doi: 10.1038/s41467-021-25800-3

Fang, C., Chen, L. Y., Nan, H. Y., Kong, L. P., Li, Y., Zhang, H. Y., et al. (2019). Rapid identification of consistent novel QTLs underlying long-juvenile trait in soybean by multiple genetic populations and genotyping-by-sequencing. *Mol. Breeding.* 39, 80. doi: 10.1007/s11032-019-0979-2

Fang, C., Liu, J., Zhang, T., Su, T., Li, S. C., Cheng, Q., et al. (2021). A recent retrotransposon insertion of *J* caused *E6* locus facilitating soybean adaptation into low latitude. *J. Integr. Plant Biol.* 63, 995–1003. doi: 10.1111/jipb.13034

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., et al. (2013). MutMap plus: genetic mapping and mutant identification without crossing in rice. *PloS One* 8, e68529. doi: 10.1371/journal.pone.0068529

Giovannoni, J. J., Wing, R. A., Ganal, M. W., and Tanksley, S. D. (1991). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Res.* 19, 6553–6558. doi: 10.1093/nar/19.23.6553

Graham, P. H., and Vance, C. P. (2003). Legumes: importance and constraints to greater use. *Plant Physiol.* 131, 872–877. doi: 10.1104/pp.017004

Han, Y., Zhao, X., Liu, D., Li, Y., Lightfoot, D. A., Yang, Z., et al. (2016). Domestication footprints anchor genomic regions of agronomic importance in soybeans. *New Phytol.* 209, 871–884. doi: 10.1111/nph.13626

Jiang, B. J., Zhang, S. W., Song, W. W., Khan, M. A. A., Sun, S., Zhang, C. S., et al. (2019). Natural variations of *FT* family genes in soybean varieties covering a wide range of maturity groups. *BMC Genomics* 20, 230. doi: 10.1186/s12864-019-5577-5

Kong, F., Liu, B., Xia, Z., Sato, S., Kim, B. M., Watanabe, S., et al. (2010). Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. *Plant Physiol.* 154, 1220–1231. doi: 10.1104/pp.110.160796

Kong, F., Nan, H., Cao, D., Li, Y., Wu, F., Wang, J., et al. (2014). A new dominant gene *E9* conditions early flowering and maturity in soybean. *Crop Sci.* 54, 2529–2535. doi: 10.2135/cropsci2014.03.0228

Langmead, B., and Salzberg, S. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, X. M., Fang, C., Xu, M. L., Zhang, F. G., Lu, S. J., Nan, H. Y., et al. (2017). Quantitative trait locus mapping of soybean maturity gene E6. *Crop Sci.* 57, 2547–2554. doi: 10.2135/cropsci2017.02.0106

Li, X. M., Fang, C., Yang, Y. Q., Lv, T. X., Su, T., Chen, L. Y., et al. (2021). Overcoming the genetic compensation response of soybean florigens to improve adaptation and yield at low latitudes. *Curr. Biol.* 31, 3755–3767. doi: 10.1016/j.cub.2021.06.037

Lin, X. Y., Fang, C., Liu, B. H., and Kong, F. J. (2021a). Natural variation and artificial selection of photoperiodic flowering genes and their applications in crop adaptation. *aBIOTECH.* 2, 156–169. doi: 10.1007/s42994-021-00039-0

Lin, X. Y., Liu, B. H., Abe, J., Weller, J. L., and Kong, F. J. (2021b). Molecular mechanisms for the photoperiodic regulation of flowering in soybean. *J. Integr. Plant Biol.* 63, 981–994. doi: 10.1111/jipb.13021

Liu, W., Jiang, B. J., Ma, L. M., Zhang, S. W., Zhai, H., Xu, X., et al. (2018). Functional diversification of flowering locus T homologs in soybean: GmFT1a and GmFT2a/5a have opposite roles in controlling flowering and maturation. *New Phytol.* 217, 1335–1345. doi: 10.1111/nph.14884

Liu, B., Kanazawa, A., Matsumura, H., Takahashi, R., Harada, K., and Abe, J. (2008). Genetic redundancy in soybean photoresponses associated with duplication of the phytochrome a gene. *Genetics.* 180, 995–1007. doi: 10.1534/genetics.108.092742

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

Lu, S. J., Dong, L. D., Fang, C., Liu, S. L., Kong, L. P., Cheng, Q., et al. (2020). Stepwise selection on homologous PRR genes controlling flowering and maturity during soybean domestication. *Nat. Genet.* 52, 428–436. doi: 10.1038/s41588-020-0604-7

Lu, S. J., Li, Y., Wang, J. L., Srinives, P., Nan, H. Y., Cao, D., et al. (2015). QTL mapping for flowering time in different latitude in soybean. *Euphytica.* 206, 725–736. doi: 10.1007/s10681-015-1501-5

Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49, 773–779. doi: 10.1038/ng.3819

Mansur, L. M., Orf, J., and Lark, K. G. (1993). Determining the linkage of quantitative trait loci to RFLP markers using extreme phenotypes of recombinant inbreds of soybean (*Glycine max* l. merr.). *Theor. Appl. Genet.* 86, 914–918. doi: 10.1007/BF00211041

Ma, X. L., Zhang, Q. Y., Zhu, Q. L., Liu, W., Chen, Y., Qiu, R., et al. (2015). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* 8, 1274–1284. doi: 10.1016/j.molp.2015.04.007

McBlain, B. A., and Bernard, R. L. (1987). A new gene affecting the time of flowering and maturity in soybeans. *J. Hered.* 78, 160–162. doi: 10.1093/oxfordjournals.jhered.a110349

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Michelmore, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis - a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U. S. A.* 88, 9828–9832. doi: 10.1073/pnas.88.21.9828

Nan, H., Cao, D., Zhang, D., Li, Y., Lu, S., Tang, L., et al. (2014). GmFT2a and GmFT5a redundantly and differentially regulate flowering through interaction with and upregulation of the bZIP transcription factor GmFDL19 in soybean. *PloS One* 9, e97669. doi: 10.1371/journal.pone.0097669

Neumaier, N., and James, A. T. (1993). Exploiting the long juvenile trait to improve adaptation of soybeans to the tropics. *ACTAR Food Legume Newsletter.* 18, 12–14.

Parvez, A. Q., and Gardner, F. P. (1987). Daylength and sowing date responses of soybean lines with "juvenile" trait. *Crop Sci.* 27, 305–310. doi: 10.2135/cropsci1987.0011183X002700020037x

Ray, J. D., Hinson, K., Mankono, J. E. B., and Malo, M. F. (1995). Genetic control of a long-juvenile trait in soybean. *Crop Sci.* 35, 1001–1006. doi: 10.2135/cropsci1995.0011183x003500040012x

Samanfar, B., Molnar, S. J., Charette, M., Schoenrock, A., Dehne, F., Golshani, A., et al. (2017). Mapping and identification of a potential candidate gene for a novel maturity locus, *E10*, in soybean. *Theor. Appl. Genet.* 130, 377–390. doi: 10.1007/s00122-016-2819-7

Schneeberger, K., and Weigel, D. (2011). Fast-forward genetics enabled by new sequencing technologies. *Trends Plant Sci.* 16, 282–288. doi: 10.1016/j.tplants.2011.02.006

Sinclair, T. R., and Hinson, K. (1992). Soybean flowering in response to the long-juvenile trait. *Crop Sci.* 32, 1242–1248. doi: 10.2135/cropsci1992.0011183X003200050036x

Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., et al. (2013a). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* 74, 174–183. doi: 10.1111/tpj.12105

Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A., Mitsuoka, C., et al. (2013). MutMap-gap: whole- genome resequencing of mutant F2 progeny bulk combined with *de novo* assembly of gap regions identifies the rice blast resistance gene pii. *New Phytol.* 200, 276–283. doi: 10.1111/nph.12369

Van Ooijen, J. W. (2004). *MapQTL5, software for the mapping of quantitative trait loci in experimental populations* (Wageningen, Netherlands: Kyazma, B. V.).

Voorrips, R. E. (2002). MapChart: software for graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77

Wang, J., Chu, S., Zhang, H., Zhu, Y., Cheng, H., and Yu., D. (2016). Development and application of a novel genome-wide SNP array reveals domestication history in soybean. *Sci. Rep.* 6, 20728. doi: 10.1038/srep20728

Wang, L. S., Fang, C., Liu, J., Zhang, T., Kou, K., Su, T., et al. (2020). Identification of major QTLs for flowering and maturity in soybean by genotyping-by-sequencing analysis. *Mol. Breeding.* 40, 99. doi: 10.1007/s11032-020-01178-w

Wang, F., Nan, H., Chen, L., Fang, C., Zhang, H., Su, T., et al. (2019). A new dominant locus, *E11*, controls early flowering time and maturity in soybean. *Mol. Breeding.* 39, 70. doi: 10.1007/s11032-019-0978-3

Watanabe, S., Hideshima, R., Xia, Z., Tsubokura, Y., Sato, S., Nakamoto, Y., et al. (2009). Map-based cloning of the gene associated with the soybean maturity locus *E3*. *Genetics.* 182, 1251–1262. doi: 10.1534/genetics.108.098772

Watanabe, S., Xia, Z. J., Hideshima, R., Tsubokura, Y., Sato, S., Yamanaka, N., et al. (2011). A map-based cloning strategy employing a residual heterozygous line reveals that the *GIGANTEA* gene is involved in soybean maturity and flowering. *Genetics.* 188, 395–407. doi: 10.1534/genetics.110.125062

Whipple, C. J., Kebrom, T. H., Weber, A. L., Yang, F., Hall, D., Meeley, R., et al. (2011). *Grassy tillers1* promotes apical dominance in maize and responds to shade signals in the grasses. *Proc. Natl. Acad. Sci. U. S. A.* 108, E506–E512. doi: 10.1073/pnas.1102819108

Xia, Z., Watanabe, S., Yamada, T., Tsubokura, Y., Nakashima, H., Zhai, H., et al. (2012). Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. *Proc. Natl. Acad. Sci. U. S. A.* 109, E2155–E2164. doi: 10.1073/pnas.1117982109

Yi, B., Chen, Y. N., Lei, S. L., Tu, J. X., and Fu, T. D. (2006). Fine mapping of the recessive genic male-sterile gene (*Bnms1*) in *Brassica napus* l. *Theor. Appl. Genet.* 113, 643–650. doi: 10.1007/s00122-006-0328-9

Yue, L., Li, X. M., Fang, C., Chen, L. Y., Yang, H., Yang, J., et al. (2021). FT5a interferes with the Dt1-AP1 feedback loop to control flowering time and shoot determinacy in soybean. *J. Integr. Plant Biol.* 63, 1004–1020. doi: 10.1111/jipb.13070

Zhao, C., Takeshima, R., Zhu, J., Xu, M., Sato, M., Watanabe, S., et al. (2016). A recessive allele for delayed flowering at the soybean maturity locus E9 is a leaky allele of FT2a, a FLOWERING LOCUS T ortholog. *BMC Plant Biol.* 16, 20. doi: 10.1186/s12870-016-0704-9

Zhong, J. S., and Kong, F. J. (2022). The control of compound inflorescences: insights from grasses and legumes. *Trends Plant Sci.* 27, 564–576. doi: 10.1016/j.tplants.2021.12.002

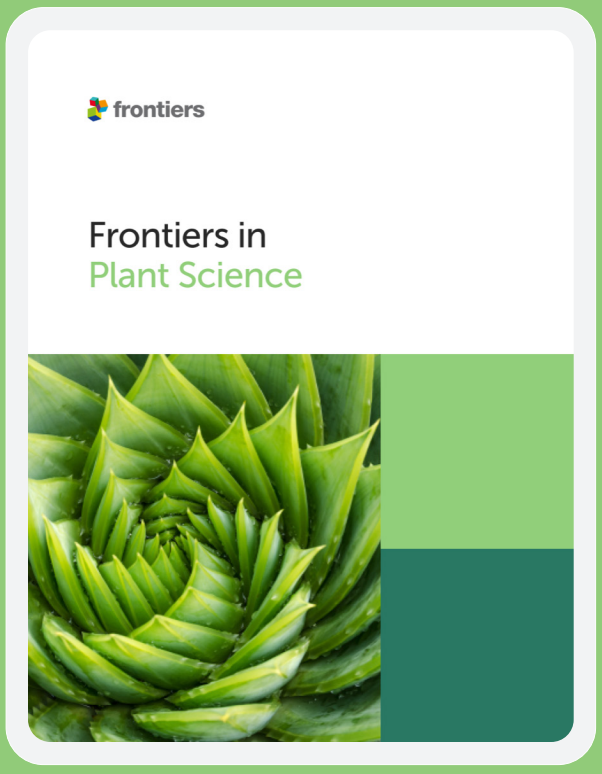
Frontiers in Plant Science

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