EPIGENETICS IN PLANT DEVELOPMENT

EDITED BY: Mingli Xu, Ullas Pedmale, Dazhong Dave Zhao, Jie Song, Gang Wu and Nobutoshi Yamaguchi

PUBLISHED IN: Frontiers in Plant Science

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-727-6
DOI 10.3389/978-2-88974-727-6

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

EPIGENETICS IN PLANT DEVELOPMENT

Topic Editors:
Mingli Xu, University of South Carolina, United States
Ullas Pedmale, Cold Spring Harbor Laboratory, United States
Dazhong Dave Zhao, University of Wisconsin–Milwaukee, United States
Jie Song, Imperial College London, United Kingdom
Gang Wu, Zhejiang Agriculture and Forestry University, China
Nobutoshi Yamaguchi, Nara Institute of Science and Technology (NAIST), Japan

Citation: Xu, M., Pedmale, U., Zhao, D. D., Song, J., Wu, G., Yamaguchi, N., eds. (2022). Epigenetics in Plant Development. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88974-727-6

Table of Contents

frontiers in Plant Science

EDITORIAL
published: 28 February 2022
doi: 10.3389/fpls.2022.864945

Editorial: Epigenetics in Plant Development

*Nobutoshi Yamaguchi**

Division of Biological Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

Keywords: chromatin structure, DNA methylation, development, epigenetics, environmental response, histone modification

Editorial on the Research Topic

Epigenetics in Plant Development

Plant growth and development are determined by the spatiotemporal regulation of gene expression and epigenetic regulators help fine-tune the timing and patterns of gene expression. For example, as a part of this Research Topic on Epigenetics in Plant Development, Yamaguchi describes recent findings about one of the best-characterized plant transcription factors, LEAFY (LFY), in *Arabidopsis*. Although many researchers have examined LFY function over the past 30 years, two independent research groups recently revealed that LFY functions as a pioneer transcription factor, one of the master regulators located at the top of the gene regulatory hierarchy. Pioneer transcription factors reprogram the closed chromatin of their target genes and thus play critical roles in specifying when and where downstream targets are expressed to ensure proper cell fate and differentiation. LFY directly binds condensed chromatin, displaces the linker histone H1 in the nucleosome, interacts with chromatin remodeling factors, and opens up chromatin to enable the binding of other factors to specify floral fate (Weigel et al., 1992; Jin et al., 2021; Lai et al., 2021). The emerging research on LFY is just one example of recent major breakthroughs in this field, but much more remains to be learned about the epigenetic mechanisms underlying plant development.

Epigenetic regulation involves multiple mechanisms, including histone modifications. Fang et al. highlight the role of histone H3 lysine methylation in regulating gene expression, with extra emphasis on reproductive development in *Arabidopsis*. Members of the SET Domain Group (SDG) serve as "writers" by depositing methylation marks (Pontvianne et al., 2010). Histone marks are recognized by "readers," such as proteins with PHD domains, WD40 repeats, and Chromo domains (Jiang et al., 2009). By contrast, LYSINE-SPECIFIC DEMETHYLASE 1 (LSD1) and Jumonji-C domain-containing proteins (JMJs) remove methylation marks, thus serving as "erasers." Yamaguchi focuses on a group of *Arabidopsis* JMJ proteins that remove trimethylation of histone H3 lysine 27 (H3K27me3). The H3K27me3 demethylases identified to date include EARLY FLOWERING 6 (ELF6)/JMJ11, RELATIVE OF ELF6 (REF6)/JMJ12, JMJ13, JMJ30, and JMJ32. These proteins often function in a redundant manner to regulate plant development and environmental responses. Keyzor et al., studied the relationship between ELF6 and JMJ13 and revealed their antagonistic functions during *Arabidopsis* flower development. Compared to the wild type, *elf6* displays increased self-fertility, whereas *jmj13* mutants show decreased self-fertility. Based on transcription data, ELF6 promotes carpel elongation by activating expansin genes. JMJ13 represses carpel growth by activating jasmonic acid signal transduction and promotes stamen growth by activating *SAUR26* expression.

Each epigenetic factor can play multiple roles in controlling gene expression in a tissue-specific manner. Ornelas-Ayala et al., introduce multiple interacting partners of ULTRAPETALA1 (ULT1) in *Arabidopsis*. ULT1 controls histone H3 lysine 4 (H3K4me3) levels and counteracts

Edited and reviewed by:
José Manuel Pérez-Pérez, Miguel Hernández University of Elche, Spain

***Correspondence:**
Nobutoshi Yamaguchi
nobuy@bs.naist.jp

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

Received: *29 January 2022*
Accepted: *07 February 2022*
Published: *28 February 2022*

Citation:
Yamaguchi N (2022) Editorial: Epigenetics in Plant Development. Front. Plant Sci. 13:864945. doi: 10.3389/fpls.2022.864945

the activity of H3K27me3 "writers." ULT1 physically interacts with the H3K4me3 writer ATX1 to induce H3K4me3 deposition, and it interacts with tissue-specific transcription factors. ULT and the GARP family transcription factor KANADI1 (KAN1) form a complex that controls gynoecium axis development. ULT and the MYB domain-containing transcription factor ULTRAPETALA INTERACTING FACTOR 1 (UIF1) control floral meristem determinacy by repressing the expression of the stem cell fate gene *WUSCHEL*. Chromatin structure is altered by chromatin remodelers, such as ATP-dependent chromatin remodeling SWITCH/SUCROSE NON-FERMENTING (SWI/SNF) complexes. The SWI/SNF complex component SWI3B was initially identified as a flowering time regulator (Sarnowski et al., 2002). Lin et al., identified a new interacting partner for SWI3B. SWI3B genetically and physically interacts with LEAF AND FLOWER RELATED (LFR) to determine adaxial-abaxial cell fate in leaves.

In addition to our knowledge of interacting partners, the factors that function upstream and downstream of each epigenetic factor are not fully understood. Jiang and Zheng summarize the current understanding of the relationship between SPOROCYTELESS/NOZZLE (the core transcription factor required for megaspore mother cell development) and epigenetic regulation at multiple layers. Hirai et al., explored factors downstream of histone deacetylase activity during xylem vessel cell differentiation and identified *OVATE FAMILY PROTEIN1* (*OFP1*), *OFP4*, and *MYB75* as downstream targets. These genes encode transcription factors that form a complex with BEL1-LIKE HOMEODOMAIN6 to control gene expression for cell differentiation.

Although the majority of reviews and research articles in this Research Topic describe work in *Arabidopsis* due to the relative ease in performing epigenetic analysis in this plant, a few researchers have performed epigenetic studies in other plant species. Zhang et al., obtained genome-wide H3K27me3 and H3K4me3 profiles in allotetraploid cotton (*Gossypium hirsutum*). In general, H3K4me3 and H3K27me3 are located around the transcription start sites of active genes and the gene bodies of silenced genes, respectively. Consistent with this notion, the presence of H3K4me3 and H3K27me3 leads to the activation and repression of gene expression, respectively, in allotetraploid cotton. Examining the roles of histone-modifying enzymes in other plant species remains an exciting area for future research.

Many studies related to this Research Topic have revealed the importance of epigenetic regulation in cell fate switching or developmental transitions. These processes occur in a limited number of cells during a limited time window. However, techniques such as chromatin immunoprecipitation followed by sequencing (ChIP-seq) require large numbers of cells and take several days to perform. To address these problems, Ouyang et al., developed an alternative method for ChIP-seq called nucleus CUT&Tag (nCUT&Tag). nCUT&Tag can be completed within a day using only 0.01 g of plant tissue as the starting material. Cao et al., explored DNA methylation dynamics using a tissue culture system to prepare plant materials at different stages of development. The combination of such sophisticated systems and highly sensitive techniques will allow researchers to further explore the epigenetic regulation of gene expression during plant development in the future.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work was supported by a grant from the Japan Science and Technology Agency PREST(JPMJPR15QA), a JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (No. 18H04782), a JSPS KAKENHI Grant-in-Aid for Scientific Research B (No. 18H02465), a Grant-in-Aid for challenging Exploratory Research (No. 19K22431), and a grant from the SECOM Science and Technology Foundation to NY.

ACKNOWLEDGMENTS

The author thanks Sachi Ando for critical comments on this manuscript.

REFERENCES

Jiang, D., Gu, X., and He, Y. (2009). Establishment of the winter-annual growth habit via *FRIGIDA*-mediated histone methylation at *FLOWERING LOCUS C* in *Arabidopsis*. *Plant Cell* 21, 1733–1746. doi: 10.1105/tpc.109.067967

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

Lai, X., Blanc-Mathieu, R., Grand Vuillemin, L., Huang, Y., Stigliani, A., Lucas, J., et al. (2021). The LEAFY floral regulator displays pioneer transcription factor properties. *Mol. Plant* 14, 829–837. doi: 10.1016/j.molp.2021.03.004

Pontvianne, F., Blevins, T., and Pikaard, C. S. (2010). Arabidopsis histone lysine methyltransferases. *Adv. Bot. Res.* 53, 1–22. doi: 10.1016/S0065-2296(10)53001-5

Sarnowski, T. J., Swiezewski, S., Pawlikowska, K., Kaczanowski, S., and Jerzmanowski, A. (2002). AtSWI3B, an Arabidopsis homolog of SWI3, a core subunit of yeast Swi/Snf chromatin remodeling complex, interacts with FCA, a regulator of flowering time. *Nucleic Acids Res.* 30, 3412–3421. doi: 10.1093/nar/gkf458

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* 69, 843–859. doi: 10.1016/0092-8674(92)90295-N

Conflict of Interest: The author declares 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: 12 January 2021
doi: 10.3389/fpls.2020.612055

Functional Identification of *EjGIF1* in *Arabidopsis* and Preliminary Analysis of Its Regulatory Mechanisms in the Formation of Triploid Loquat Leaf Heterosis

*Chao Liu[1], Renwei Huang[2], Lingli Wang[3] and Guolu Liang[4]**

[1]College of Basic Medical Sciences, Chengdu University of Traditional Chinese Medicine, Chengdu, China, [2]Sichuan Provincial Key Laboratory for Development and Utilization of Characteristic Horticultural Biological Resources, College of Chemistry and Life Sciences, Chengdu Normal University, Chengdu, China, [3]Technical Advice Station of Economic Crop, Chongqing, China, [4]College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Shunquan Lin,
South China Agricultural University, China
Yuan Wang,
University of California, Riverside, United States

***Correspondence:**
Guolu Liang
lianggl@swu.edu.cn

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

Received: *30 September 2020*
Accepted: *11 December 2020*
Published: *12 January 2021*

Citation:
Liu C, Huang R, Wang L and Liang G (2021) Functional Identification of EjGIF1 in Arabidopsis and Preliminary Analysis of Its Regulatory Mechanisms in the Formation of Triploid Loquat Leaf Heterosis. Front. Plant Sci. 11:612055. doi: 10.3389/fpls.2020.612055

Although several results have been obtained in triploid loquat heterosis (i.e., leaf size of triploid loquat) studies in the past years, the underlying mechanisms of the heterosis are still largely unknown, especially the regulation effects of one specific gene on the corresponding morphology heterosis. In this study, we sought to further illustrate the regulatory mechanisms of one specific gene on the leaf size heterosis of triploid loquats. A leaf size development-related gene (*EjGIF1*) and its promoter were successfully cloned. Ectopic expression of *EjGIF1* in *Arabidopsis* showed that the leaf size of transgenic plantlets was larger than that of WTs, and the transgenic plantlets had more leaves than WTs. Quantitative Reverse Transcription PCR (qRT-PCR) showed that the expression level of *EjGIF1* showed an AHP expression pattern in most of the hybrids, and this was consistent with our previous phenotype observations. Structure analysis of *EjGIF1* promoter showed that there were significantly more light-responsive elements than other elements. To further ascertain the regulatory mechanisms of *EjGIF1* on triploid loquat heterosis, the methylation levels of *EjGIF1* promoter in different ploidy loquats were analyzed by using bisulfite sequencing. Surprisingly, the total methylation levels of *EjGIF1* promoter in triploid showed a decreasing trend compared with the mid-parent value (MPV), and this was also consistent with the qRT-PCR results of *EjGIF1*. Taken together, our results suggested that *EjGIF1* played an important role in promoting leaf size development of loquat, and demethylation of *EjGIF1* promoter in triploid loquats caused *EjGIF1* to exhibit over-dominance expression pattern and then further to promote leaf heterosis formation. In conclusion, *EjGIF1* played an important role in the formation of triploid loquat leaf size heterosis.

Keywords: triploid loquat, *EjGIF1*, transgenic *Arabidopsis*, leaf size, heterosis, DNA demethylation

INTRODUCTION

Heterosis, or hybrid vigor, is a common phenomenon in many diploid or polyploid organisms, which means the biomass, resistance ability, yield, and some other agronomic traits in hybrids are greater than that of the parents (Hofmann, 2012). Heterosis has been widely used to improve the yield of the field crops and vegetables continuously and thus has greatly solved the crisis of food shortage especially in some developing countries (Agbo and Teixeira da Silva, 2014). However, to date we still know little about the mechanisms of heterosis (Wang et al., 2015). Researchers have proposed several models from the genetic aspect to explain the mechanisms of heterosis including dominance, over-dominance, and epistasis, but none of these models can fully explain this phenomenon (Jones, 1917; East, 1936; Yu et al., 1997). Recent studies on maize, soybean, rice, *Arabidopsis*, etc., have found that heterosis may be associated with the differential gene expression based on the fact that no new genes are produced after hybridization (Guo and Rafalski, 2013; Miller et al., 2015; Wang et al., 2015; Taliercio et al., 2017; Chen et al., 2018). Two gene expression-related models, additive and non-additive gene expression, were proposed by Chen (2010) to further explain heterosis phenomenon. With the development of functional genomics, such as the application of RNA-Seq technology, more and more studies have found that heterosis may be highly related to additive expression pattern due to the fact that genes exhibit non-additive expression pattern in hybrids are comparatively rare, and the non-additive genes are deemed to associate with the formation of transgressive traits in hybrids (Guo et al., 2006; Thiemann et al., 2014). For instance, study on triploid loquat, Liu et al. (2018a) analyzed the leaf transcriptomes of the triploid loquats and their parents in two cross combinations and identified that 94.56 and 86.97% transcripts were expressed additively in the two cross combinations, respectively, and only 5.44 and 13.03% genes expressed non-additively. These results indicated that additively expressed genes may play a fundamental role in the formation of triploid loquats.

Recent studies found that epigenetic mechanisms, especially DNA methylation which are considered to be associated with the regulation of gene expression in a number of plant species (Arikan et al., 2018). Due to the regulatory function on gene expression, DNA methylation level is also considered to be associated tightly with heterosis (Nakamura and Hosaka, 2010). Studies have shown that DNA methylation is mainly occurred in the CpG island of the promoter, and the DNA methylation density of a promoter can affect the transcriptional activity of the gene (De Smet et al., 1999; Alasaari et al., 2012).

Loquat [*Eriobotrya japonica* (Thunb.) L.; 2n = 2x = 34] belongs to the subtribe Pyrinae in the Rosaceae family and is favored by many people due to its excellent flavor and medicinal applications (Wu et al., 2015). However, the loquat fruits sold in the market currently are all diploid with too many seeds, and this significantly affects their edibility (Liu et al., 2018a). Triploid loquat breeding provides a new way to solve the problem of low edible rate of diploid loquats. Previous studies in our lab found that triploid loquats are not only seedless, but also have a variety of excellent traits that diploid and tetraploid loquats do not have, such as larger and greener loquat leaves, showing an obvious heterosis (Liu et al., 2018b, 2019). Liu et al. (2018a,b) have studied the mechanisms of triploid loquat heterosis by using several triploid loquats with clear genetic relationship and found that extensive genetic variation and DNA methylation remodeling after the formation of triploid loquat may change the gene expression patterns in triploid loquats, and these further promoted the formation of triploid loquat heterosis. However, for triploid loquat heterosis, we still know little about the mechanisms.

Leaves are the photosynthetic place of plants, absorbing sunlight energy to synthesize biological energy (Gonzalez et al., 2012; Jiao et al., 2019). Leaves of eudicots are initiated at the flank of the shoot apical meristem (SAM), and the extent and direction of leaf growth have a great influence on the leaf size and shape (Horiguchi et al., 2005; Vercruyssen et al., 2015). Plant Growth-Regulating Factor (GRFs) is a family of transcription factors that regulate leaf development, and nine GRFs (GRF1-GRF9) were identified from *Arabidopsis* (Kim et al., 2003). Studies on *Arabidopsis* and rice found that GRFs could repress or activate the expression of their target genes by binding to the regulatory region of DNA (Kim et al., 2012; Kuijt et al., 2014). Overexpression of *AtGRF1*, *AtGRF2*, and *AtGRF5* could lead the cell number or size of transgenic leaves to decrease, and these make the transgenic plants have larger leaves than wild-type (WT) plants (Kim et al., 2003). GRF INTERACTING FACTOR 1/ANGUSTIFOLIA 3 (GIF1/AN3) is a transcriptional coactivator which is a functional homolog to the human synovial sarcoma translocation protein (SYT) transcription coactivator (Horiguchi et al., 2005; Vercruyssen et al., 2015). Overexpression of *AtGIF1* enlarged the leaf size of the transgenic plants, whereas, loss-of-function *gif1* plants developed narrower leaves (Kim and Kende, 2004; Horiguchi et al., 2005). Yeast two-hybrid analysis showed that AtGIF1 could interact with both AtGRF1 and AtGRF5, and positively promoted the leaf cell proliferation and regulated the leaf size in plants (Kim and Kende, 2004; Horiguchi et al., 2005). Thus, like GRFs, GIF1 also functions as an important transcription factor in the size and shape regulation of plant leaves (Kim and Kende, 2004).

Although we have verified that the triploid loquat leaves become larger, greener than that of diploid and tetraploid loquats, showing an obvious heterosis (**Supplementary Material**: **Supplementary Table S1**), we still know little about the association of leaf development with triploid loquat leaf heterosis and also few reports on this issue. Illuminating the mechanisms of leaf development of loquat could help us better understand the heterosis phenomenon of triploid loquat leaf and provide more details for the triploid loquat application in loquat breeding. In this study, we have identified the transcription factor EjGIF1 in loquat and made a further validation for *EjGIF1* function, and at the same time, *EjGIF1* promoter was cloned and also the methylation level of *EjGIF1* promoter was analyzed by bisulfite sequencing (BSP) in different ploidy loquats. Our study will provide more information on the morphology heterosis of triploid loquat leaf.

MATERIALS AND METHODS

Plant Lines

In order to overcome the unclear origin of loquat, the triploid loquats used in this study were created by cross-fertilizing. Two triploid loquat lines were generated in 2003, named Triploid-A and Triploid-B. For the two triploid lines, the same female parent (Longquan-1 tetraploid) was used to cross with two different wild diploid loquats, GC-1 (Triploid-A) and GC-23 (Triploid-B). The tetraploid parent Longquan-1 was selected by our laboratory, while the wild diploid parents, GC-1 and GC-23 were identified in Guizhou Province, China, which grow naturally in the rocky arid region and have strong levels of abiotic and biotic resistance (Wu et al., 2015). In the meantime, GC-1 and GC-23 also have a far genetic distance with cultivated loquats, which could increase mutations in triploid loquats after hybridization (Wu et al., 2015). Finally, nine and three triploid loquats were obtained in Triploid-A and Triploid-B, respectively, which were labeled as A-1, A-2, A-3,… A-9 and B-1, B-2, B-3. All the plants were grown in a natural environment, in the Experimental Base of College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China.

Isolation of *EjGIF1* Complementary DNA Sequence

The reference sequence of *EjGIF1* was obtained from the RNA-Seq data base in our laboratory. The leaf material of Longquan-1 tetraploid was used for the cDNA isolation; moreover, the RNA extraction and cDNA synthesis methods were performed the same as Liu et al. (2019). The cloning primers (*EjGIF1-Asc* I-F and *EjGIF1-Xba* I-R) were designed based on the *EjGIF1* reference sequence, and the restriction enzyme sites were added at the 5'-end and 3'-end for the subsequent vector construction (**Table 1**). PCR products were then cloned to the pMD19-T (Takara, Dalina) for sequencing.

Isolation and Analysis of *EjGIF1* Promoter Sequence

In order to analyze the structure and methylation level of *EjGIF1 Cis*-element, the promoter sequence of *EjGIF1* was isolated based on the user manual of Universal Genome-Walker Kit 2.0 (Takara, Clontech Laboratories, Inc., Japan). The nested primers (1-*EjGIF1* GSP1 to 5-*EjGIF1* GSP2) used for promoter cloning were listed in **Table 1**, and the amplification products were sequenced as the same as described above. The possible regulatory elements of the *EjGIF1* promoter were annotated by using the PlantCARE database.

Expression Pattern Analysis of Loquat *EjGIF1* Gene in Different Ploidy and Developmental Stages

To analyze the expression level of *EjGIF1* in different ploidy loquats, and in different developmental stages of loquat leaves as well, leaves from three developmental periods of different ploidy loquats were collected and named P I (young leaves < 5 cm), P II (5 cm < medium mature leaves < 15 cm) P III (mature leaves), respectively (Gong et al., 2014). The expression levels in different developmental stages were analyzed by using the materials of P I, P II, and P III. The RNAs were extracted as described by Liu et al. (2019). cDNA synthesis and Quantitative Reverse Transcription-PCR (qRT-PCR) methods were also performed by using the methods as described by Liu et al. (2019). The primers (q*EjGIF1*-F and q*EjGIF1*-R) used in qRT-PCR analysis were listed in **Table 1**. *Actin* of loquat was analyzed with the primer sequences 5'-ATCCTTCGTCTGGACCTTGC-3' and 5'-GACAATTTCCCGT TCAGCAGT-3'. All of the samples were examined in triplicate.

EjGIF1 Overexpression Plasmid Construction and *Arabidopsis* Transformation

The full length cDNA sequence of *EjGIF1* was cloned to *Asc* I-*Xba* I sites of pFGC5941 plasmid so that the *EjGIF1* could express under the control of CaMV 35S promoter. The recombinant plasmid pFGC5941-35S::*EjGIF1* was then transferred to the *Agrobacterium tumefaciens* strain LBA4404 by means of electric shock. Afterward, WT plants were transformed by using the floral dip method (Clough and Bent, 1998) with minor modifications. Infiltration media used contained 5% sucrose and 0.02% Silwet. Seeds of transgenic lines (T0) were planted in soil and were selected by spraying with 20 g/L glufosinate-ammonium after 2 weeks. The same selection methods were used until the T2 generation was obtained, and the T2 homozygous progenies were used for phenotype observation and expression test of *EjGIF1* by qRT-PCR. All the seedlings with glufosinate-ammonium resistance were grown in a growth chamber under the 16 h light/8 h dark photoperiod (2,500 lux).

Positive Transgenic Plantlet Verification

Genomic DNAs were isolated from young, fresh leaves of glufosinate-ammonium resistance plants and WT plants with a modified cetyltrimethyl ammonium bromide (CTAB) method (Liu et al., 2005). Then, PCR was carried out for detecting the insertion, and the WT was used as a control. The transgenic and WT plants were tested for the presence of both *EjGIF1* and CaMV 35s genes separately, and primers (*EjGIF1*-F and *EjGIF1*-R, and CaMV 35s_F and CaMV 35s_R) are listed in **Table 1**.

Gene Expression Detection in the Positive and Wild Type Plants

Total RNAs were extracted from young, fresh leaves of T2 homozygous progenies and WT plants. The RNA extraction, cDNA synthesis and qRT-PCR methods were performed the same as Liu et al. (2019). The primers (q*EjGIF1*-F and q*EjGIF1*-R) were listed in **Table 1**. WTs were used as controls and the reference gene (*Actin*) of *Arabidopsis* was analyzed with the primer sequences 5'-CTTCGTCTTCCACTTCAG-3' and 5'-ATC ATACCAGTCTCAACAC-3'. Each transgenic line and each WT was examined in three plantlets as biological repetition.

TABLE 1 | Primers used in this study.

Primer name	Primer sequence (5'-3')	Tm value (°C)
EjGIF1-Asc I-F	5'-AGGCGCGCCATGCAGCAGCACCTGATCAGA-3'	87.3
EjGIF1-Xba I-R	5'-GCTCTAGATTAATTTCCATCATCGGTCGAT-3'	68.6
1-*EjGIF1* GSP1	5'-TGGTAGGAGGCTGGGGTTGAGAATC-3'	68.7
1-*EjGIF1* GSP2	5'-GCTGTAGCTTTGCTTGGTTCTCTGC-3'	66.0
2-*EjGIF1* GSP1	5'-CTCTCTCTAACTTTCTCACTCC-3'	49.6
2-*EjGIF1* GSP2	5'-GCTTTTTTTTTACAGAGTTGAG-3'	51.9
3-*EjGIF1* GSP1	5'-TTGCTGCATGTAATGTGCTCCTGGTTG-3'	71.1
3-*EjGIF1* GSP2	5'-AGATTCCGCTGTAGCTTTGCTTGGTTC-3'	69.4
4-*EjGIF1* GSP1	5'-AAGAAGGAGGACCTGCTGAATGTGATC-3'	67.4
4-*EjGIF1* GSP2	5'-GTTGTTAGGATAATAGGCTGCCATCAT-3'	63.8
5-*EjGIF1* GSP1	5'-CAGATTGTTGAGATGTTTATTGCGGGC-3'	69.1
5-*EjGIF1* GSP2	5'-AATGGCGTACAGAGAATGCGATTGTCA-3'	69.9
q*EjGIF1*-F	5'-TACTCCCAGCAACCGTTTTCA-3'	60.7
q*EjGIF1*-R	5′-TCCAGCATTATTTCCCTCATT-3'	56.7
EjGIF1-F	5'-ATGCAGCAGCACCTGATG-3'	55.1
EjGIF1-R	5'-TTAATTTCCATCATCGGTCGAT-3'	51.5
CaMV 35s_F	5'-TGAGACTTTTCAACAAAGGATAATT-3'	54.6
CaMV 35s_R	5'-TGTCCTCTCCAAATGAAATGAAC-3'	58.5
CpG1-F	5'-ACAGTTACCTGAGGACTCTGGAGTC-3'	64.4
CpG1-R	5'-CTGTGGTAGTGAGAAGTAAGGTCGT-3'	65.2
CpG2-F	5'-CCACAGTAAGTACAACCACCAG-3'	59.6
CpG2-R	5'-CTCAAACAGATCGTGTCTACACTTT-3'	58.5
CpG3-F	5'-GGTTTTGTAGGTAAGATTATAGATTTGAGA-3'	61.6
CpG3-R	5'-TAAAAATAATCCCCAACCACCTATA-3'	59.4

Leaf Morphology Traits Analysis of Transgenic and WT Plants

The T2 homozygous progenies and WT plants were grown in the growth chamber for about 1 month and their leaf morphology traits were recorded individually. The methods for leaf length and width measuring were the same as Liu et al. (2018b), and the leaf size was measured by using the ImageJ software. For each trait, 20 individuals in each transgenic line were measured as biological repetition, and three values were measured for each individual as technical repetition.

BSP Sequencing for *EjGIF1* Promoter in Different Ploidy Loquats

Tiangen Bisulfite Conversion Kit (Tiangen Company, Beijing) was adopted for genomic DNA bisulfite conversion. CpG islands prediction and PCR amplification primers for bisulfite sequencing design were carried out by using the online software (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi). Primers (CpG1-F and CpG1-R, CpG2-F and CpG2-R, and CpG3-F and CpG3-R) used for bisulfite sequencing are listed in **Table 1**. The amplification products were also sequenced as the same as described above, and for each CpG island, 15 randomly chosen clones per genotype were sequenced. The methylation levels were counted as described by Liu et al. (2018b).

Statistical Analysis

The phylogenetic tree was generated by using the Clustal W, and the bootstrap test was set at 1,000 to test confidence for the tree (Higgins et al., 1994). The MEGA 5.0 software was used for phylogenetic tree construction with Neighbor-Joining method (Tamura et al., 2011). Mid-parent value (MPV) was adopted to measure the heterosis, and it was calculated by using the method of Turner (1953). Briefly, MPV was calculated according to the genomic contribution by the two parents, i.e., MPV = 2/3 Longquan-1tetraploid + 1/3 GC-1/GC-23. The gene expression patterns were classified into two classes by using the method described by Liu et al. (2019). Briefly, (i) additive expression pattern, which gene expression levels in hybrids were at the MPV (MPL); (ii) non-additive expression pattern, which gene expression level was deviated from the MPV. The non-additive expression pattern was further classified into two classes; (iii) dominance expression pattern, which the gene expression level was at the high parent level (HPL) or at the low parent level (LPL); and (iv) over-dominance expression pattern, which the gene expression level was above the high parent level (AHP) or below the low parent level (BLP). Finally, the significance examination was performed by using the one-way ANOVA method.

RESULTS

Identification and Characterization of *EjGIF1*

Based on the reference sequence from RNA-Seq database, a segment of 651 bp cDNA sequence was obtained and sequenced, named *EjGIF1*. Sequence analysis showed that *EjGIF1* encoded

a 216 amino acids protein with the molecular weight of 23.20 kDa. The sequence of *EjGIF1* was submitted to National Center for Biotechnology Information (NCBI) and the accession number was MK573556. To investigate the relationship among the *GIF1* genes in different species, we downloaded the reported cDNA sequences from NCBI, and these reported GIF1 proteins were mainly distributed in 10 families, Rosaceae, Solanaceae, Curcurbitaceae, Euphorbiaceae, Malvaceae, Sterculiaceae, Leguminosae, Rutaceae, Juglandaceae, and Papilionoideae. The phylogenetic tree was created by using the deduced protein of EjGIF1 and these reported GIF1 proteins. The same with the traditional taxonomy, our results showed that EjGIF1 was separated from the other GIF1 proteins clearly, and EjGIF1 was clustered into Rosaceae and was closest to *Malus* × *domestica* (**Figure 1B**).

GIF1 is a leaf shape related protein which was first isolated by Relichova (1976). Previous researches showed that GIF1 is a homolog of SYT whose N-terminal contains a conserved SYT N-terminal homology (SNH) domain, and this domain could participate in protein-protein interactions (Crew et al., 1995; Thaete et al., 1999; Kato et al., 2002). In this study, results of multiple sequence alignment showed that EjGIF1 also contained a SNH domain, and this was consistent with the previous studies (**Figure 1A**). Taken together, these results suggested that EjGIF1 gene is kept highly conserved during the evolution processes.

Generating and Verification of Transformants

To investigate the potential function of *EjGIF1*, an over-expression vector with *EjGIF1* CDS sequence under the control of CaMV 35S promoter was transferred into *Arabidopsis*. After continuous screening with glufosinate-ammonium, we finally got 50 plantlets belonging to 10 transgenic lines. The transgenic seedlings and the WT ones were then transferred to the new pots and cultured in the growth chamber.

To verify the reliability of the transgenic plantlets, the presence of *EjGIF1* and CaMV_35s in the genomes of transgenic and WT plantlets were performed by PCR separately. The empty vector (pFGC5941) and the WT genomic DNA were set as controls. The detection results of the two genes in the transformants and WTs suggested that the two expected specific fragments appeared in the right positions, indicating the precision of these transgenic plantlets was reliable (**Figure 2**).

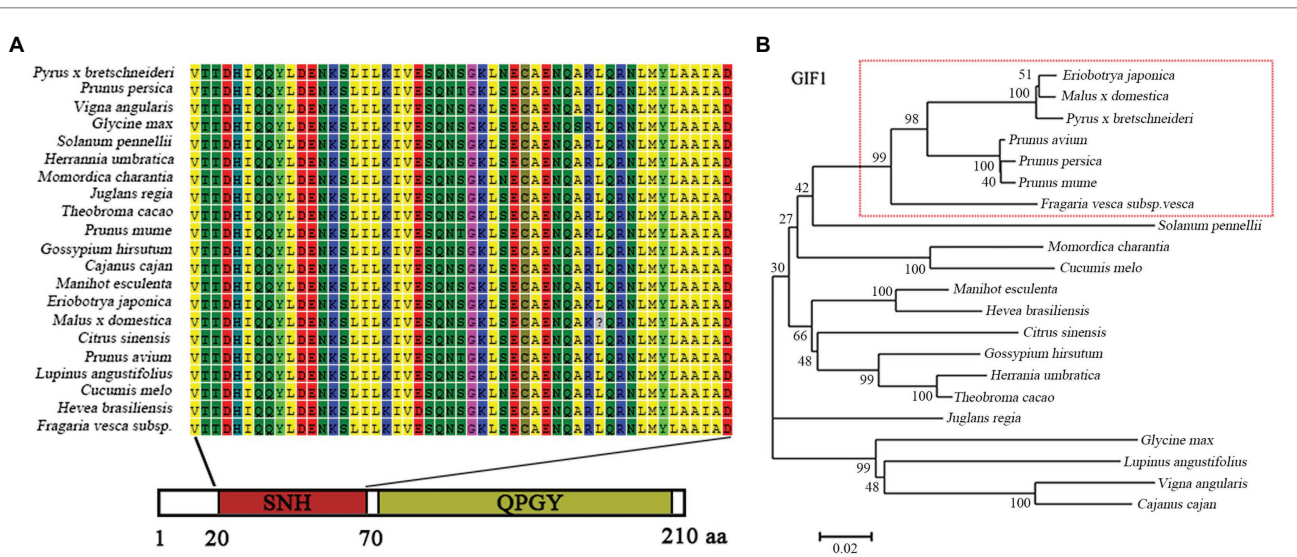

FIGURE 1 | (A) Multiple sequence alignment of EjGIF1 SNH domains with other GIF1s reported in National Center for Biotechnology Information (NCBI); **(B)** Phylogenetic relationships between loquat EjGIF1 and other GIF1 proteins reported in NCBI. The unrooted were constructed using MEGA 5.0 by neighbor-joining method, and a bootstrap test was set at 1,000 to test confidence for the tree.

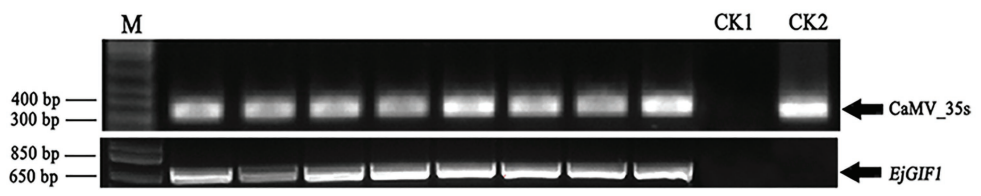

FIGURE 2 | Positive transgenic plantlets verification. CaMV_35s and *EjGIF1* genes were detected by PCR, and CK1 and CK2 were set as control. M represents Marker; CK1: genome DNA of wild type; CK2: empty pFGC5941 plasmid.

Over-Expression of *EjGIF1* in *Arabidopsis* Enlarged Leaf Size and Leaf Number

To evaluate the regulatory effects of *EjGIF1* on leaf development in *Arabidopsis*, we selected five independent transgenic lines (named: OE1, OE2…OE5) with fine phonetype for further phenotype analysis, and the WT plantlets were set as controls. The transgenic and WT plantlets were grown in a growth chamber for about 1 month. The same to the previous studies, we also found that overexpression *EjGIF1* in *Arabidopsis* could lead the leaf length and width to become larger than the WTs (**Table 2**). Moreover, the leaf sizes of the transgenic plantlets were enlarged as well. As shown in **Table 2**, the leaf area of the five transgenic plantlets (OE1, OE2…OE5) were 342.53 mm^2, 380.32 mm^2, 313.00 mm^2, 285.72 mm^2, and 285.48 mm^2, while the WT was 257.71 mm^2 (**Figures 3A,C**; **Table 2**). Correlation analysis between the leaf area and the expression level of *EjGIF1* in the five transgenic plantlets found that except for OE1, there was a positive correlation between the leaf area and the expression level of *EjGIF1* in the transgenic plantlets (**Figures 3B,C**). Therefore, these indicated that *EjGIF1* plays an important role in regulating the development of the loquat leaf size.

Interestingly, we also found that the transgenic plantlets had significantly more leaves than the WT (**Figure 3D**; **Table 2**). As shown in **Table 2**, the WT contains 11 leaves, while the transgenic plantlets contain 20, 20, 19, 15, and 23 leaves, respectively. Different from previous studies (Kim and Kende, 2004; Horiguchi et al., 2005), our results suggested that *EjGIF1* could not only promote the development of leaf size but also increase the formation of leaf primordium, but how does this occur requires to be further researched.

Finally, the expression levels of *EjGIF1* in the five transgenic lines and WTs were detected by qRT-PCR. Results showed that *EjGIF1* were expressed higher in all the five transgenic lines than that of the WTs (**Figure 3B**), and transcripts have not been detected out in the WT ones.

Expression Analysis of *EjGIF1* in Different Developmental Stages and Ploidy Loquat

To ascertain the expression levels of *EjGIF1* in different developmental stages of loquat leaf, we then measured the expression levels of *EjGIF1* in three developmental stages of different ploidy loquats by qRT-PCR. Our results showed that, for most of the genotypes, the expression levels of *EjGIF1* displayed a tendency of rising first and then dropping, and expressed the highest levels in P II (**Figure 4A**).

Our previous studies on the morphologies of loquat leaves demonstrated that many morphological characteristics of triploid loquat leaves showed a different degree of heterosis compared with their parents, such as leaf length and width (**Supplementary Material**: **Supplementary Table S1**). In order to investigate the regulatory effects of *EjGIF1* on the formation of triploid leaf morphology heterosis, the expression analyses of *EjGIF1* in different ploidy loquats were performed. Based on the results above, materials of P II were used for further analysis. The results showed that the expression of *EjGIF1* in most of the hybrids exhibited AHP (A-3, A-4, A-5, A-6, A-7, A-8, and B-2) expression pattern, demonstrating pronounced heterosis (**Figure 4B**; **Table 3**). Only A-1 showed an LPL expression pattern, and A-2, A-9, B-1, and B-3 were expressed BLP (**Figure 4B**; **Table 3**). No hybrids expressed MPL and HPL. The qRT-PCR results were basically consistent with our previous morphology (leaf length and width) studies. These results indicated that *EjGIF1* may play an important role in the formation of leaf heterosis of triploid loquat.

Isolation and Characterization of *EjGIF1* Promoter

Gene expression was regulated by both *Cis*-elements and *trans*-regulatory factors (Shi et al., 2012; Wittkopp and Kalay, 2012). In order to ascertain the possible regulatory mechanisms of *EjGIF1* gene in regulating the leaf development of loquats, we cloned a 2,475 bp promoter sequence from the upstream of the initiation codon of *EjGIF1* by using the Longquan-1 tetraploid genomic DNA. Results of the online prediction showed that there were five hormone-responsive elements (GARE-motif, TATC-box, TCA-element, ABRE, and P-box), 12 light-responsive elements (AE-box, Box4, C-box, G-box, GAG-motif, Gap-box, LAMP-element, Sp1, TCT-motif, CATT-motif, I-box, and MNF1), and six stress-responsive elements (HSE, ARE, GC-motif, MBS, DRE, and TC-rich repeats; **Table 4**; **Figure 5A**). What caught our attention was that the light-responsive elements were far more than the other elements, and these suggested that the expression of *EjGIF1* may be highly sensitive to light changes, but this need to be further validated.

Promoter Methylation Level Analysis of Different Ploidy Loquat

DNA methylation level of a promoter can directly affect the transcriptional activity of the gene (Wei et al., 2018). Moreover, gene expression level could further affect the

TABLE 2 | Leaf morphologies analysis of the transgenic and WT plantlets.

	WT	OE1	OE2	OE3	OE4	OE5
Leaf length (cm)	3.3 ± 0.1[a]	4.3 ± 0.1	4.5 ± 0.1	4.3 ± 0.1	4.0 ± 0.1	4.7 ± 0.1
Leaf width (cm)	1.2 ± 0.1	1.3 ± 0.1	1.6 ± 0.1	1.4 ± 0.1	1.4 ± 0.1	1.3 ± 0.1
leaf area (mm^2)	257.71 ± 1.36	342.53 ± 4.20	380.32 ± 0.91	313.00 ± 1.55	285.72 ± 3.32	285.48 ± 1.01
Leaf number	11 ± 1	20 ± 1	20 ± 1	19 ± 1	15 ± 2	23 ± 2

[a]*mean ± standard deviation.*

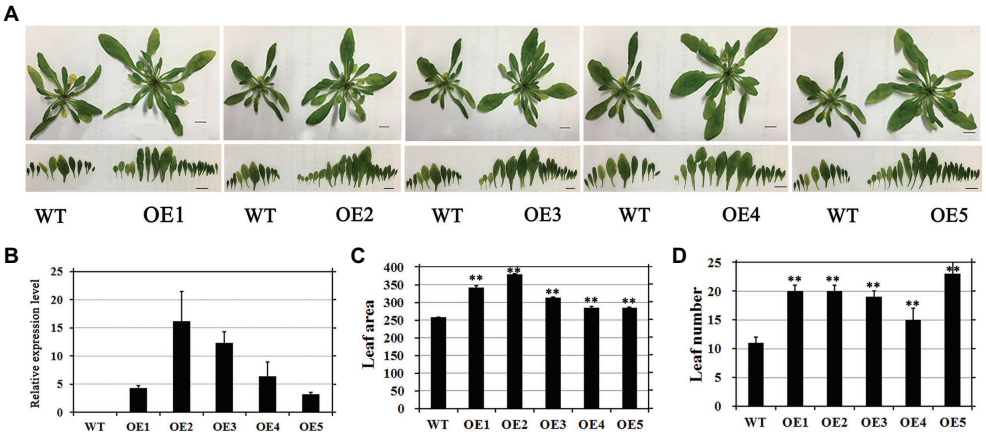

FIGURE 3 | (A) Phenotypes of transformants and WT plants. The size of the bar showed in the picture was 1 cm. **(B)** Expression analysis of *EjGIF1* in T2 homozygous progenies and WTs. *Actin* gene was selected by our laboratory previously which was used as a control. All data are from three biological repeats (n = 3). **(C)** Leaf area analysis of the transgenic plantlets and WTs. **(D)** Leaf number analysis of the transgenic plantlets and WTs. **Represents the significance level of the one-way ANOVA test, p = 0.01. Error bars denote |S|D.

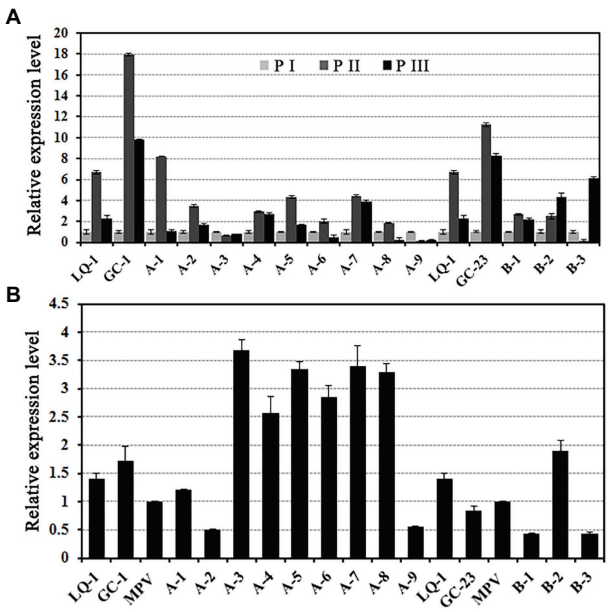

FIGURE 4 | (A) *EjGIF1* expression analyses in different development stages and **(B)** ploidy loquats. *Actin* gene was selected by our laboratory previously which was used as a control. All data are from three technical repeats (n = 3). Error bars denote |S|D.

TABLE 3 | *EjGIF1* expression patterns in the hybrids.

	Additive	Dominance		Over-dominance	
	MPL[a]	HPL[b]	LPL[c]	AHP[d]	BLP[e]
EjGIF1	NONE	NONE	A-1	A-3, A-4, A-5, A-6, A-7, A-8, B-2	A-2, A-9, B-1, B-3

[a]*Gene expression level was at the MPV.*
[b]*Gene expression level was at the high parent level.*
[c]*Gene expression level was at the low parent level.*
[d]*Gene expression level is above the high parent level.*
[e]*Gene expression level is below the low parent level.*

phenotype of a plant. Therefore, in order to further analyze the regulatory effects of *EjGIF1* on the leaf morphology development of loquat, we randomly selected three triploids from Triploid-A (A-3, A-5, and A-6), and the methylation levels of *EjGIF1* promoters in the three triploids and their parents (Longquan-1 tetraploid, GC-1) were analyzed by bisulfite sequencing. CpG island prediction showed that there were three CpG islands in the promoter, and the length were 151 bp, 306 bp, and 191 bp, respectively (**Figure 5B**). The sequences of the three CpG islands were further used for primer design (**Table 1**). Bisulfite sequencing results exhibited that methylation levels of diploid parent GC-1 were basically slightly higher than that of the tetraploid parent in all the three contexts (^{m}CG, ^{m}CHG, and ^{m}CHH) among the 3 CpG islands. However, when compared with MPVs, the methylation levels of the hybrids (A-3, A-5, and A-6) showed a decreasing trend in almost all the three methylation types among CpG1 and CpG3 islands, and only CpG2 showed an increasing trend (**Figures 6A–C**). Interestingly, when we counted for the total methylation level for the *EjGIF1* promoter in the three hybrids, it was showed that the methylation level demonstrated a decreasing level in all the three hybrids compared with MPV (21.50%), with the methylation level of 17.56% (A-3), 18.33% (A-5), and 17.84% (A-6), respectively (**Figure 6D**).

Taken together, our results suggested that the total methylation levels of *EjGIF1* promoter in triploid loquats (A-3, A-5, and A-6) showed a decreasing trend, and this may generate the expression differences of *EjGIF1* between triploid loquats and

TABLE 4 | Partial *Cis*-regulatory elements in the promoter of *EjGIF1*.

	Motif	Sequence	Function
EjGIF1	AE-box	AGAAACAA	Part of a module for light response
	ARE	TGGTTT	*Cis*-acting regulatory element essential for the anaerobic induction
	Box 4	ATTAAT	Part of a conserved DNA module involved in light responsiveness
	C-box	CTGACGTCAG	*Cis-acting regulatory element involved in light responsiveness*
	G-box	CACGAC	*Cis-acting regulatory element involved in light responsiveness*
	GAG-motif	AGAGAGT	Part of light responsive element
	GARE-motif	AAACAGA	Gibberellin-responsive element
	Gap-box	AAATGGAGA	Part of light responsive element
	LAMP-element	CCAAAACCA	Part of light responsive element
	MBS	CAACTG	MYB binding site involved in drought-inducibility
	Sp1	CC(G/A)CCC	Light responsive element
	TATC-box	TATCCCA	*Cis-acting element involved in gibberellin-responsiveness*
	TCA-element	CCATCTTTTT	*Cis-acting element involved in salicylic acid responsiveness*
	TCT-motif	TCTTAC	Part of light responsive element
	Circadian	CAANNNNATC	*Cis-acting regulatory element involved in circadian control*
	ABRE	TACGTG	*Cis-acting element involved in the abscisic acid responsiveness*
	C-repeat/DRE	TGGCCGAC	Regulatory element involved in cold- and dehydration responsiveness
	CATT-motif	GCATTC	Part of a light responsive element
	GC-motif	CCCCCG	Enhancer-like element involved in anoxic specific inducibility
	HSE	AAAAAATTTC	*Cis-acting element involved in heat stress responsiveness*
	I-box	GATATGG	Part of light responsive element
	MNF1	GTGCCC(A/T)	Light responsive element
	P-box	CCTTTTG	Gibberellin-responsive element
	TC-rich repeats	ATTTTCTTCA	*Cis*-acting element involved in defense and stress responsiveness

parents (Longquan-1 tetraploid, GC-1), and further regulate the leaf morphology heterosis of triploid loquat.

DISCUSSION

Leaf is an important organ of plant photosynthesis, and it can directly affect the accumulation of sugar. In the meantime, it is also an important aspect for the plant morphology formation, and determines the growth potential of plant (Yan et al., 2008). Leaf size or leaf area greatly determines the light interception and transpiration (Monteith, 1977). Researches on leaf development have been lasted for many years. In previous studies, many transcription factors, such as GRFs or AINTEGUMENTA 3 (GIF1), that regulate leaf development have been verified and reported, and also some regulatory mechanisms of these transcription factors have been validated (Gonzalez et al., 2012; Dkhar and Pareek, 2014). Kuijt et al. (2014) found that *Oskn2*, an upstream sequence of *KNOX* gene, could interact with *OsGRF3* and *OsGRF10* in rice. In *Arabidopsis*, it was found that the expression levels of *GRFs* were regulated by *miR396*, and overexpressing *miR396* could cause narrow-leaf phenotypes (Liu et al., 2009). So far, studies on leaf development are mainly focused on the model plants, grasses, or herbaceous plants, such as *Arabidopsis*, barley, *Brassica napus* etc., and there are relatively few studies on the leaf development of woody plants (Mizukami and Fischer, 2000; Osnato et al., 2010; Liu et al., 2012; Dkhar and Pareek, 2014). In this study, we have successfully cloned a transcriptional coactivator GIF1 from loquat (EjGIF1), and our phylogenetic tree analysis showed that EjGIF1 is highly homologous with plants of the Rosaceae family, and is kept highly conserved during the evolution processes. Results of *EjGIF1* function validation demonstrated that the ectopic expression of *EjGIF1* in *Arabidopsis* could increase the leaf size, and this was consistent with previous findings (Kim et al., 2002; Horiguchi et al., 2005). Interestingly, we also found that the transgenic plantlets contained more leaves than the WTs. These results suggested that *EjGIF1* may play an important role in the leaf development of the *Arabidopsis*.

Polyploid possesses more than two sets of chromosome per cell, and it plays an important role in the plant evolution (Sattler et al., 2016). Delighting, polyploidization is often accompanied with the increased growth vigor of the plants compared with the diploid progenitors, and so does the triploid loquat (Stebbins, 1971; Chen, 2007; Li et al., 2017). Despite the ploidy effect, triploid loquat demonstrated pronounced heterosis compared with the diploid and tetraploid loquats based on our previous studies on the cultivated triploid loquats (Liu et al., 2018a,b, 2019). For the mechanisms studies of triploid loquat heterosis, some results have been obtained, but the molecular mechanisms of triploid loquat heterosis are still poorly understood (Liu et al., 2018a,b, 2019). As described above, to date, researches on the correlation between heterosis and genes are mainly on the whole genome-wide expression levels, and few studies have been performed on some specific genes. In this study, we have investigated the expression level of one specific leaf development-related gene *EjGIF1* in triploid loquats and their parents based on the results of our previous research that the leaf morphologies (length and width) of triploid loquats exhibited pronounced heterosis. Based on the results of *EjGIF1* ectopic expression in *Arabidopsis*, we further investigated the expression level of *EjGIF1* in triploid loquats

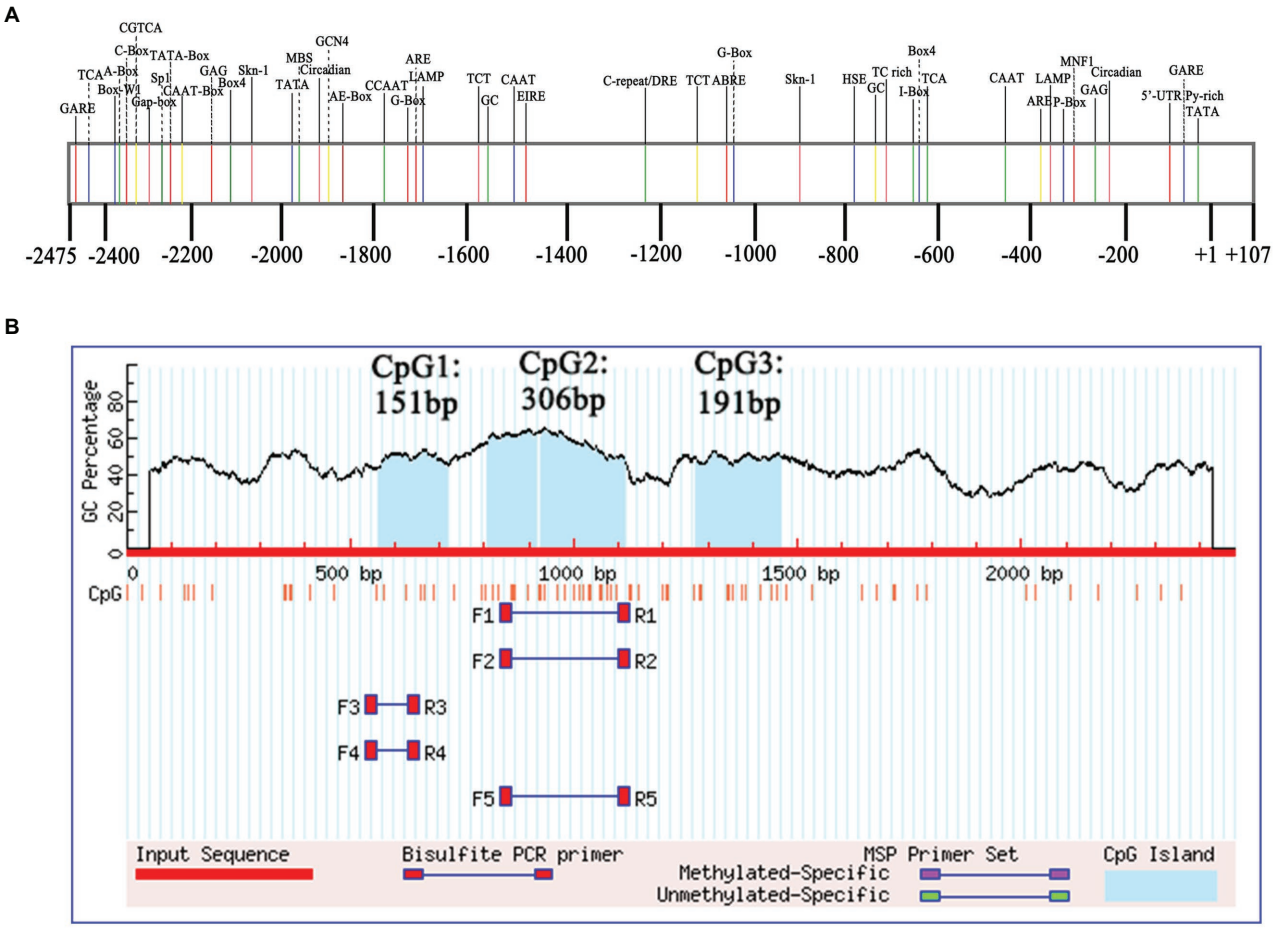

FIGURE 5 | (A) Structure analysis and **(B)** CpG islands prediction of *EjGIF1* promoter.

and their parents. It was found that *EjGIF1* was expressed AHP in most of the triploid loquats, showing a non-additive expression pattern, and this was basically consistent with our previous studies on leaf morphology heterosis of triploid loquats (Liu et al., 2018b). These suggested that high expression of *EjGIF1* in triploid hybrids played a critical role in the leaf size heterosis formation.

Gene expression was greatly regulated by *Cis*-element, which could affect the transcriptional efficiency and stability (Garí et al., 1997; Mei et al., 2008). In order to ascertain the structure of *EjGIF1* promoter, we have successfully obtained a 2,475 bp promoter sequence by using the method of genome walking. After making a prediction for the promoter online, it was found that the light-responsive elements were significantly more than other elements, suggesting that the expression of *EjGIF1* may be greatly sensitive to light changes. In fact, many studies have found that light can affect the leaf size development, for example, light quality affects the trophic effects through photosynthesis and further determines the leaf morphogenesis or leaf area (Tardieu et al., 1999; Cookson and Granier, 2006). In this study, we indeed found that there were more light-responsive elements in the *EjGIF1* promoter, so we suggested that the expression of *EjGIF1* may be largely regulated by light changes. On the other hand, Baldissera et al. (2014) studied the alfalfa plants and found that plant branch development and the number of shoot per plant were most affected by light. Furthermore, Horiguchi et al. (2005) found that *AN3* was expressed at a high level in the basal region of leaf primordia, therefore, based on the results discussed above, we further proposed that *EjGIF1* could also promote the formation of leaf primordium. If this is the case, the transgenic *Arabidopsis* of *EjGIF1* should have more leaves than the WTs. Intriguingly, the transgenic plantlets did have more leaves than the WTs. Taken together, we speculated that the expression of *EjGIF1* was greatly induced by the light changes, and *EjGIF1* may also have an effect on the formation of leaf primordia. However, whether or how the light works on these issues are important questions and still need to be deeply studied.

Recent studies found that polyploidization could trigger extensive DNA methylation remodeling in the first or the following few generations due to the fact that it is an effective

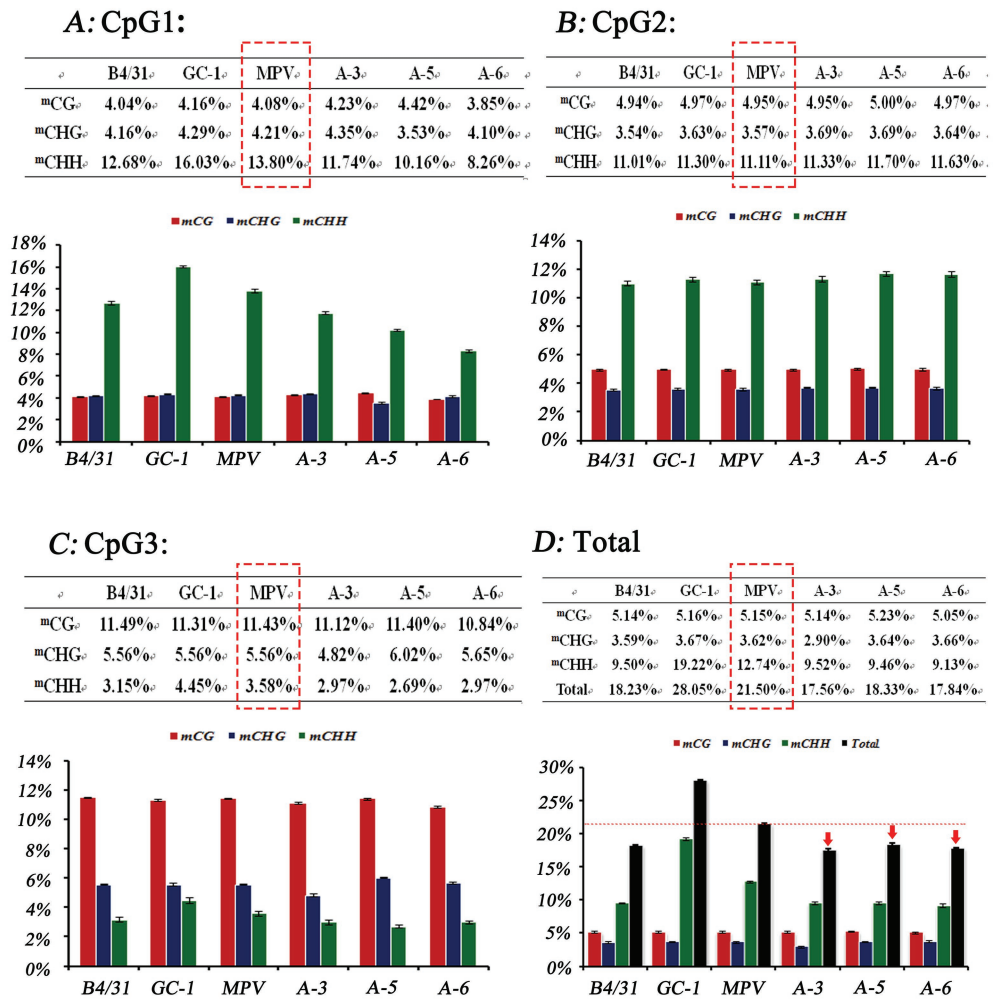

FIGURE 6 | Cytosine methylation level analysis of the three CpG islands using bisulfite sequencing. Each CpG islands was sequenced by using 15 PCR clones. The collective methylation levels (%) of the three types of cytosine residues, ^{m}CG, ^{m}CHG, and ^{m}CHH for the three CpG islands (CpG1, CpG2, and CpG3) were depicted in **(A–C)**, and the total methylation level of the promoter was depicted in **(D)**. Error bars denote |S|D.

way for polyploid to maintain the genome stability (Wang et al., 2004; Fort et al., 2016). That is, DNA methylation occurred in the whole genome could regulate gene expression, inhibit transposable elements (TEs) transposition, and maintain the structure stability of chromatin (Feinberg, 2007; Bucher et al., 2012). Among them, methylation through promoter region is an effective way to inhibit gene expression without DNA sequence variation (Maunakea et al., 2010). These make more and more researchers believe that there must be a correlation between heterosis and DNA methylation, and begin to explain the heterosis mechanisms from the aspect of epigenetic (Groszmann et al., 2013; Xiong et al., 2013; Wang et al., 2018). To ascertain the methylation level of *EjGIF1* promoter among triploid loquats and their parents, we have analyzed the methylation levels in three randomly selected triploid loquats and their parents by bisulfite sequencing. It was found that the total methylation level of *EjGIF1* promoter in triploid loquats showed a decreasing trend compared with MPV, and this was consistent with the qRT-PCR results. Since the three hybrids (A-3, A-5, and A-6) used for methylation level analysis were selected randomly in this study, it was worth noting that the expression levels of *EjGIF1* in some hybrids exhibited a low expression level, and we still did not know the methylation levels of *EjGIF1* promoter in these hybrids. Therefore, the methylation levels of these low expressed hybrids need to be further detected for verifying the association between the expression level and the methylation level.

Taken together, our results suggested that (1) compared with previous studies, our study found that *EjGIF1* showed significant regulation effects on the development of leaf size; and (2) demethylation of *EjGIF1* promoter made *EjGIF1* exhibit over-dominance expression pattern in triploid loquats, and this further promoted the formation of triploid loquat heterosis. In short, *EjGIF1* played an important role in the formation of triploid loquat leaf size heterosis.

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

CL and RH conducted the experiments and the paper writing. LW helped to revise the English. GL provided the funding support. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by "Xinglin Scholar" Scientific Research Promotion Program for Academic Talents of Chengdu University of Traditional Chinese Medicine (BSH2019018).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.612055/full#supplementary-material

REFERENCES

Agbo, C. U., and Teixeira da Silva, J. A. (2014). Expression of heterosis and heritability in vegetative traits of *Gongronema latifolia*. *Open J. Genet.* 4, 146–156. doi: 10.4236/ojgen.2014.42015

Alasaari, J., Lagus, M., Ollila, H. M., Toivola, A., Kivimäki, M., Vahtera, J., et al. (2012). Environmental stress affects DNA methylation of a CpG rich promoter region of serotonin transporter gene in a nurse cohort. *PLoS One* 7:e45813. doi: 10.1371/journal.pone.0045813

Arikan, B., Özden, S., and Turgut-Kara, N. (2018). DNA methylation related gene expression and morphophysiological response to abiotic stresses in *Arabidopsis thaliana*. *Environ. Exp. Bot.* 149, 17–26. doi: 10.1016/j.envexpbot.2018.01.011

Baldissera, T. C., Frak, E., Carvalho, P. C., and Louarn, G. (2014). Plant development controls leaf area expansion in alfalfa plants competing for light. *Ann. Bot.* 113, 145–157. doi: 10.1093/aob/mct251

Bucher, E., Reinders, J., and Mirouze, M. (2012). Epigenetic control of transposon transcription and mobility in *Arabidopsis*. *Curr. Opin. Plant Biol.* 15, 503–510. doi: 10.1016/j.pbi.2012.08.006

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annu. Rev. Plant Biol.* 58, 377–406. doi: 10.1146/annurev.arplant.58.032806.103835

Chen, Z. J. (2010). Molecular mechanisms of polyploid and hybrid vigor. *Trends Plant Sci.* 15, 57–71. doi: 10.1016/j.tplants.2009.12.003

Chen, L., Bian, J., Shi, S., Yu, J., Khanzada, H., Wassan, G. M., et al. (2018). Genetic analysis for the grain number heterosis of a super-hybrid rice WFYT025 combination using RNA-Seq. *Rice* 11:37. doi: 10.1186/s12284-018-0229-y

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.

Cookson, S. J., and Granier, C. (2006). A dynamic analysis of the shade-induced plasticity in *Arabidopsis thaliana* rosette leaf development reveals new components of the shade-adaptative response. *Ann. Bot.* 97, 443–452. doi: 10.1093/aob/mcj047

Crew, A. J., Clark, J., Fisher, C., Gill, S., Grimer, R., Chand, A., et al. (1995). Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to the Kruppel-associated box in human synovial sarcoma. *EMBO J.* 14, 2333–2340.

De Smet, C., Lurquin, C., Lethé, B., Martelange, V., and Boon, T. (1999). DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. *Mol. Cell. Biol.* 19, 7327–7335.

Dkhar, J., and Pareek, A. (2014). What determines a leaf's shape? *EvoDevo* 5:47. doi: 10.1186/2041-9139-5-47

East, E. M. (1936). Heterosis. *Genetics* 21, 375–397.

Feinberg, A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature* 447, 433–440. doi: 10.1038/nature05919

Fort, A., Ryder, P., McKeown, P. C., Wijnen, C., Aarts, M. G., Sulpice, R., et al. (2016). Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in *Arabidopsis thaliana*. *New Phytol.* 209, 590–599. doi: 10.1111/nph.13650

Garí, E., Piedrafita, L., Aldea, M., and Herrero, E. (1997). A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* 13, 837–848.

Gong, W., Qi, P., Du, J., Sun, X., Wu, X., Song, C., et al. (2014). Transcriptome analysis of shade-induced inhibition on leaf size in relay intercropped soybean. *PLoS One* 9:e98465. doi: 10.1371/journal.pone.0098465

Gonzalez, N., Vanhaeren, H., and Inze, D. (2012). Leaf size control: complex coordination of cell division and expansion. *Trends Plant Sci.* 17, 332–340. doi: 10.1016/j.tplants.2012.02.003

Groszmann, M., Greaves, I. K., Fujimoto, R., Peacock, W. J., and Dennis, E. S. (2013). The role of epigenetics in hybrid vigour. *Trends Genet.* 29, 684–690. doi: 10.1016/j.tig.2013.07.004

Guo, M., and Rafalski, J. A. (2013). "Gene expression and heterosis in maize hybrids" in *Polyploid and hybrid genomics*. eds. Z. J. Chen and J. A. Birchler (New York: Wiley), 59–84.

Guo, M., Rupe, M. A., Yang, X., Crasta, O., Zinselmeier, C., Smith, O. S., et al. (2006). Genome-wide transcript analysis of maize hybrids: allelic additive gene expression and yield heterosis. *Theor. Appl. Genet.* 113, 831–845. doi: 10.1007/s00122-006-0335-x

Higgins, D., Thompson, J., Gibson, T., Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.

Hofmann, N. R. (2012). A global view of hybrid vigor: DNA methylation, small RNAs, and gene expression. *Plant Cell* 24:841. doi: 10.1105/tpc.112.240312

Horiguchi, G., Kim, G. T., and Tsukaya, H. (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of *Arabidopsis thaliana*. *Plant J.* 43, 68–78. doi: 10.1111/j.1365-313X.2005.02429.x

Jiao, K., Li, X., Guo, Y., Guan, Y., Guo, W., Luo, D., et al. (2019). Regulation of compound leaf development in mungbean (*Vigna radiata* L.) by CUP-SHAPED COTYLEDON/NO APICAL MERISTEM (CUC/NAM) gene. *Planta* 249, 765–774. doi: 10.1007/s00425-018-3038-z

Jones, D. F. (1917). Dominance of linked factors as a means of accounting for heterosis. *Genetics* 3, 310–312.

Kato, H., Tjernberg, A., Zhang, W., Krutchinsky, A. N., An, W., Takeuchi, T., et al. (2002). SYT associates with human SNF/SWI complexes and the C-terminal region of its fusion partner SSX1 targets histones. *J. Biol. Chem.* 277, 5498–5505. doi: 10.1074/jbc.M108702200

Kim, J. H., Choi, D. S., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. *Plant J.* 36, 94–104. doi: 10.1046/j.1365-313x.2003.01862.x

Kim, J. H., and Kende, H. (2004). A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13374–13379. doi: 10.1073/pnas.0405450101

Kim, J. S., Mizoi, J., Kidokoro, S., Maruyama, K., Nakajima, J., Nakashima, K., et al. (2012). *Arabidopsis* growth-regulating factor7 functions as a transcriptional

repressor of abscisic acid- and osmotic stress-responsive genes, including DREB2A. *Plant Cell* 24, 3393–3405. doi: 10.1105/tpc.112.100933

Kim, G. T., Shoda, K., Tsuge, T., Cho, K. H., Uchimiya, H., Yokoyama, R., et al. (2002). The ANGUSTIFOLIA gene of *Arabidopsis*, a plant CtBP gene, regulates leaf-cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell wall formation. *EMBO J.* 21, 1267–1279. doi: 10.1093/emboj/21.6.1267

Kuijt, S. J., Greco, R., Agalou, A., Shao, J., Hoen, C. C., Overnas, E., et al. (2014). Interaction between the GROWTH-REGULATING FACTOR and KNOTTED1- LIKE HOMEOBOX families of transcription factors. *Plant Physiol.* 164, 1952–1966. doi: 10.1104/pp.113.222836

Li, X., Shahid, M. Q., Xia, J., Lu, Z., Fang, N., Wang, L., et al. (2017). Analysis of small RNAs revealed differential expressions during pollen and embryo sac development in autotetraploid rice. *BMC Genomics* 18:129. doi: 10.1186/s12864-017-3526-8

Liu, J., Hua, W., Yang, H. L., Zhan, G. M., Li, R. J., Deng, L. B., et al. (2012). The *BnGRF2* gene (GRF2-like gene from *Brassica napus*) enhances seed oil production through regulating cell number and plant photosynthesis. *J. Exp. Bot.* 63, 3727–3740. doi: 10.1093/jxb/ers066

Liu, C., Liu, T., Ohlson, E. W., Wang, L., Wu, D., Guo, Q., et al. (2019). Loquat *Eriobotrya japonica* (Thunb.) circadian clock gene cloning and heterosis studies of artifcial triploid loquat. *Sci. Hortic.* 246, 328–337. doi: 10.1016/j.scienta.2018.10.068

Liu, D., Song, Y., Chen, Z., and Yu, D. (2009). Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in *Arabidopsis*. *Physiol. Plant.* 136, 223–236. doi: 10.1111/j.1399-3054.2009.01229.x

Liu, C., Wang, M., Wang, L., Guo, Q., and Liang, G. (2018b). Extensive genetic and DNA methylation variation contribute to heterosis in triploid loquat hybrids. *Genome* 61, 437–447. doi: 10.1139/gen-2017-0232

Liu, C., Wu, D., Wang, L., Dang, J., He, Q., Guo, Q., et al. (2018a). *Cis*-regulated additively expressed genes play a fundamental role in the formation of triploid loquat (*Eriobotrya japonica* (Thunb.) Lindl.) Heterosis. *Mol. Gen. Genomics* 293, 967–981. doi: 10.1007/s00438-018-1433-6

Liu, Y. X., Yang, X. H., Lin, S. Q., Hu, G. B., and Liu, C. M. (2005). An improved procedure for nuclear DNA isolation from Eriobotrya plants and its application. *J. Fruit Sci.* 22, 182–185. doi: 10.3969/j.issn.1009-9980.2005.02.021

Maunakea, A. K., Nagarajan, R. P., Bilenky, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466, 253–257. doi: 10.1038/nature09165

Mei, G., Sean, Y., Mary, R., Bin, H., David, R. B., Lane, A., et al. (2008). Genome-wide allele-specific expression analysis using massively parallel signature sequencing (MPSSTM) reveals *cis*- and *trans*-effects on gene expression in maize hybrid meristem tissue. *Plant Mol. Biol.* 66, 551–563. doi: 10.1007/s11103-008-9290-z

Miller, M., Song, Q., Shi, X., Juenger, T. E., and Chen, Z. J. (2015). Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of *Arabidopsis*. *Nat. Commun.* 6:7453. doi: 10.1038/ncomms8453

Mizukami, Y., and Fischer, R. L. (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 97, 942–947. doi: 10.1073/pnas.97.2.942

Monteith, J. L. (1977). Climate and the efficiency of crop production in Britain. *Philos. Trans. R. Soc. B: Biol. Sci.* 281, 277–294.

Nakamura, S., and Hosaka, K. (2010). DNA methylation in diploid inbred lines of potatoes and its possible role in the regulation of heterosis. *Theor. Appl. Genet.* 120, 205–214. doi: 10.1007/s00122-009-1058-6

Osnato, M., Stile, M. R., Wang, Y., Meynard, D., Curiale, S., Guiderdoni, E., et al. (2010). Cross talk between the KNOX and ethylene pathways is mediated by intron-binding transcription factors in barley. *Plant Physiol.* 154, 1616–1632. doi: 10.1104/pp.110.161984

Relichova, J. (1976). Some new mutants. *Arabidopsis Inf. Serv.* 13, 25–28.

Sattler, M. C., Carvalho, C. R., and Clarindo, W. R. (2016). The polyploidy and its key role in plant breeding. *Planta* 243, 281–296. doi: 10.1007/s00425-015-2450-x

Shi, X., Ng, D. W. K., Zhang, C., Comai, L., Ye, W., and Chen, Z. J. (2012). *Cis*- and *trans*-regulatory divergence between progenitor species determines gene-expression novelty in *Arabidopsis* allopolyploids. *Nat. Commun.* 3:950. doi: 10.1038/ncomms1954

Stebbins, G. L. (1971). *Chromosomal evolution in higher plants*. London: Addison-Wesley.

Taliercio, E., Eickholt, D., Rouf, R., and Carter, T. (2017). Changes in gene expression between a soybean F1 hybrid and its parents are associated with agronomically valuable traits. *PLoS One* 12:e0177225. doi: 10.1371/journal.pone.0177225

Tamura, K., Peterson, D., Peterson, N., Steche, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121

Tardieu, T., Granier, C., and Muller, B. (1999). Modelling leaf expansion in a fluctuating environment: are changes in specific leaf area a consequence of changes in expansion rate? *New Phytol.* 143, 33–43.

Thaete, C., Brett, D., Monaghan, P., Whitehouse, S., Rennie, G., Rayner, E., et al. (1999). Functional domains of the SYT and SYT-SSX synovial sarcoma translocation proteins and co-localization with the SNF protein BRM in the nucleus. *Hum. Mol. Genet.* 8, 585–591.

Thiemann, A., Fu, J., Seifert, F., Grant-Dwonton, R. T., Schrag, T. A., Pospisil, H., et al. (2014). Genome-wide meta-analysis of maize heterosis reveals the potential role of additive gene expression at pericentromeric loci. *BMC Plant Biol.* 14:88. doi: 10.1186/1471-2229-14-88

Turner, J. H. (1953). A study of heterosis in upland cotton I. yield of hybrids compared with varieties. II. Combining ability and inbreeding effect. *Agron. J.* 45, 485–490.

Vercruyssen, L., Tognetti, V. B., Gonzalez, N., Van Dingenen, J., De Milde, L., Bielach, A., et al. (2015). GROWTH REGULATING FACTOR5 stimulates Arabidopsis chloroplast division, photosynthesis, and leaf longevity. *Plant Physiol.* 167, 817–832. doi: 10.1104/pp.114.256180

Wang, H., Fang, Y., Wang, L., Zhu, W., Ji, H., Wang, H., et al. (2015). Heterosis and differential gene expression in hybrids and parents in *Bombyx mori* by digital gene expression profling. *Sci. Rep.* 5:8750. doi: 10.1038/srep08750

Wang, J., Tian, L., Madlung, A., Lee, H. S., Chen, M., Lee, J. J., et al. (2004). Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. *Genetics* 167, 1961–1973. doi: 10.1534/genetics.104.027896

Wang, Y., Zhang, K., Sun, L., Han, X., Fan, S., Li, X., et al. (2018). Study on the relationship between genetic variation of DNA methylation and heterosis in soybean leaves. *Euphytica* 214:85. doi: 10.1007/s10681-018-2161-z

Wei, D., Li, A., Zhao, C., Wang, H., Mei, C., Khan, R., et al. (2018). Transcriptional regulation by CpG sites methylation in the core promoter region of the bovine SIX1 gene: roles of histone H4 and E2F2. *Int. J. Mol. Sci.* 19:213. doi: 10.3390/ijms19010213

Wittkopp, P. J., and Kalay, G. (2012). *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* 13, 59–69. doi: 10.1038/nrg3095

Wu, D., Fan, W., He, Q., Guo, Q., Spano, A. J., Wang, Y., et al. (2015). Genetic diversity of loquat (*Eriobotrya japonica* (Thunb.) Lindl.) native to Guizhou Province (China) and its potential in the genetic improvement of domesticated cultivars. *Plant Mol. Biol. Report.* 33, 952–961. doi: 10.1007/s11105-014-0809-y

Xiong, W., Li, X., Fu, D., Mei, J., Li, Q., Lu, G., et al. (2013). DNA methylation alterations at 59-CCGG sites in the interspecific and intraspecific hybridizations derived from *Brassica rapa* and *B. napus*. *PLoS One* 8:e65946. doi: 10.1371/journal.pone.0065946

Yan, S., Yan, C. J., and Gu, M. H. (2008). Molecular mechanism of leaf development. *Hereditas* 30, 1127–1135. doi: 10.3724/sp.j.1005.2008.01127

Yu, S. B., Li, J. X., Tan, Y. F., Gao, Y. J., Li, X. H., Zhang, Q. F., et al. (1997). Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proc. Natl. Acad. Sci. U. S. A.* 94, 9226–9231.

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 in Plant Science

ORIGINAL RESEARCH
published: 11 February 2021
doi: 10.3389/fpls.2020.620282

The Copy Number Variation of *OsMTD1* Regulates Rice Plant Architecture

*Qing Liu[1†], Jinke Xu[1†], Yunhua Zhu[2], Yuxing Mo[1], Xue-Feng Yao[3], Ruozhong Wang[1], Wenzhen Ku[1], Zhigang Huang[1], Shitou Xia[1], Jianhua Tong[1], Chao Huang[1], Yi Su[1], Wanhuang Lin[1], Keqin Peng[1], Chun-Ming Liu[3,4] and Langtao Xiao[1]**

[1] Hunan Provincial Key Laboratory of Phytohormones and Growth Development, College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, China, [2] Hengyang Medical College, University of South China, Hengyang, China, [3] Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing, China, [4] Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China*

Copy number variation (CNV) may have phenotypic effects by altering the expression level of the gene(s) or regulatory element(s) contained. It is believed that CNVs play pivotal roles in controlling plant architecture and other traits in plant. However, the effects of CNV contributing to special traits remain largely unknown. Here we report a CNV involved in rice architecture by modulating tiller number and leaf angle. In the genome of *Oryza sativa* ssp. *japonica* cv. Nipponbare, we found a locus *Loc_Os08g34249* is derived from a 13,002-bp tandem duplication in the nearby region of *OsMTD1*, a gene regulating tillering in rice. Further survey of 230 rice cultivars showed that the duplication occurred in only 13 *japonica* rice cultivars. Phenotypic investigation indicated that this CNV region may contribute to tiller number. Moreover, we revealed that *OsMTD1* not only influences rice tiller number and leaf angle, but also represses *pri-miR156f* transcription in the CNV region. Intriguingly, this CNV performs function through both the dosage and position effects on *OsMTD1* and *pri-miR156f*. Thus, our work identified a CNV and revealed a molecular regulatory basis for its effects on plant architecture, implying this CNV may possess importance and application potential in molecular breeding in rice.

Keywords: copy number variation, *OsMTD1*, rice, plant architecture, *pri-miR156f*

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Gang Wu,
Zhejiang Agriculture and Forestry University, China
Shengben Li,
Nanjing Agricultural University, China

***Correspondence:**
Langtao Xiao
ltxiao@hunau.edu.cn

[†] These authors have contributed equally to this work

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

Received: *22 October 2020*
Accepted: *31 December 2020*
Published: *11 February 2021*

Citation:
Liu Q, Xu J, Zhu Y, Mo Y, Yao X-F, Wang R, Ku W, Huang Z, Xia S, Tong J, Huang C, Su Y, Lin W, Peng K, Liu C-M and Xiao L (2021) The Copy Number Variation of OsMTD1 Regulates Rice Plant Architecture. Front. Plant Sci. 11:620282. doi: 10.3389/fpls.2020.620282

INTRODUCTION

Genomic rearrangements include duplications, deletions, and inversions of unique genomic segments at specific regions, as well as translocations, marker chromosomes, isochromosomes, and other complex rearrangements (Lupski, 1998; Feuk et al., 2006; Weckselblatt and Rudd, 2015). These rearrangements are not random events, but instead the reflection of higher-order architectural features of the genome (Lee and Lupski, 2006; Żmien'ko et al., 2014). Different from the whole genome duplication in a cell, the copy number variation (CNV) is the microduplication and deletion, which means an abnormal number of copies of one or more segments of DNA (Sebat et al., 2004). A CNV is commonly regarded as a DNA segment that has been deleted, inserted, or duplicated on certain chromosomes. The length of DNA is more than 1 kb and variable in copy number in comparison with a reference genome (Feuk et al., 2006). Previous studies indicated that

CNVs not only involve in intraspecific genome variations, but also cause phenotypic differences. Thus, CNVs can be developed as markers for molecular identification. Genetic diversity can be differentiated by analyzing CNVs (Żmien'ko et al., 2014).

It was reported that CNVs are in variable linkage disequilibrium with flanking SNPs (Hinds et al., 2006; Locke et al., 2006; Yu et al., 2013). CNV could underlie a significant proportion of normal variation including differences in various features (Lee and Lupski, 2006). Known data suggest that CNV mainly affects the members of large families of functionally redundant genes, and the effects of individual CNV events on phenotype are usually modest (Żmien'ko et al., 2014). Altering copy number of a gene family member may only trigger quantitative rather than qualitative changes, making the CNV–phenotype association difficult to be detected. Increasing evidences showed that copy number polymorphisms contribute to natural genetic variation and adaptability in plants; some CNVs for specific genes have been linked to important traits such as flowering time, plant height, and stress resistance (Żmien'ko et al., 2014). A dramatic fruit size change due to a CNV with an insertion of 6–8 kb that affected gene regulation was described during tomato breeding (Cong et al., 2008). In wheat, a CNV has been found to determine the extreme dwarf phenotype by tandem segmental duplication of a region containing the green revolution gene *Rht-D1b* in the haploid genome (Li Y. et al., 2012).

Rice (*Oryza sativa*) is an important staple food crop in the world and a model plant of monocots; whether and how its CNVs are associated with specific traits have also been widely concerned. A CNV at the *GL7* locus has been reported; a tandem duplication of a 17.1-kb segment leads to an increase in grain length (Wang Y. et al., 2015). A 1,212-bp deletion of *qSW5* has been reported to be clearly associated with an increase in rice grain width (Shomura et al., 2008). It has been also reported that a natural tandem array of a 3,137-bp sequence in the upstream of *IPA1* leads to superior yielding (Zhang et al., 2017). Although the knowledge of CNVs in higher plants is still poor, recent studies confirmed the prevalence of CNVs in the *Oryza* species and suggested that CNVs probably play a far more significant role in plant development than previously thought. High-level CNVs existing in different rice cultivars might associate to phenotypic diversity, yet how they affect yield, quality, resistance, and development processes is largely unknown (Li S. et al., 2012; Yu et al., 2013).

OsMTD1 is a tillering-related gene in rice (Liu et al., 2015). Here we describe a previously unknown transcriptional mechanism that OsMTD1 is able to repress *pri-miR156f* transcripts by the position effect. Furthermore, we provide evidences showing that *OsMTD1*-located region involved a CNV, a tandem segmental duplication resulting in the increasing expression of the *OsMTD1* and reduction of tiller number. This CNV harbors a 13,002-bp region on the eighth chromosome, covering one protein-coding gene *OsMTD1* and a microRNA precursor of osa-miR156f. The results by surveying a panel of 190 rice cultivars showed that 13 of 82 *japonica* cultivars harboring two copies of CNV corresponding sequence by segmental tandem duplication produce less tillers than the one-copy normal cultivars. Transgenic experiments indicated that the *OsMTD1* not only influences tiller number and leaf angle, but also regulates *pri-miR156f* transcription in this CNV region.

MATERIALS AND METHODS

Plant Materials, Field Trails, and Tiller Number Investigation

The mini-core collection accessions from the China National Crop Gene Bank in the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, as described in **Supplementary Table 1**, were used in our experiments. Another *japonica* cv. Kitaake was used for CRISPR/Cas9 editing and overexpression analysis. In addition, tobacco (*Nicotiana benthamiana*) leaves were used for *Agrobacterium*-mediated transient expression analysis.

Rice tiller number investigations were conducted in Beijing. Different rice cultivars were transplanted to a paddy field with single plant per hill. The tiller number was counted from three to six randomly chosen individual hills at heading stage in summer of 2011 and autumn of 2013, respectively (**Supplementary Table 1**).

Sequence Alignments and Comparisons

The bacterial artificial chromosome sequences from *japonica* cv. Nipponbare and *indica* cv. 93–11 were used to determine the start or the end point range in sequence of *OsMTD1*-located CNV. Then, the 13,002-bp reference genome sequence from Nipponbare was used in BLASTN (National Center for Biotechnology Information) searches against different rice databases for other cultivars, including *japonica* cv. Zhonghua 11 and *indica* cv. Zhenshan 97, Minghui 63, 93–11, Shuhui 498, and RP Bio-226, to determine their orthologous regions. The conserving segments, InDels, and substitution mutations in the orthologous regions of *indica* and *japonica* were identified by using the BLAST, MEGA, and DNAMAN programs.

Plasmid Construction and Plant Transformation

The vector constructions for the CRISPR/Cas9-mediated gene editing were performed as previously described (Miao et al., 2013). The vectors for *OsMTD1* overexpression in which the *OsMTD1* gene was driven by the CaMV 35S promoter were constructed as previously described (Liu et al., 2015). The constructs were transformed into ZH11 or Kitaake by *Agrobacterium tumefaciens*–mediated transformation (Hiei and Komari, 2008).

Expressions in tobacco leaves were performed in two different plasmids of *pCAMBIA1301* and *pSN1301* vectors using Golden Gate cloning strategy. The *pSN1301* is an adapted form of *pCAMBIA1301* in which a CaMV 35S promoter was added. The region containing the native sequence of *OsMTD1* and *pre-miR156f* was amplified from a *japonica* cv. Nipponbare genomic DNA. The DNA fragment for *pCAM1301::MTD1-OsmiR156f* was amplified by primers 5′-gga tcc ccg ggt acc TGG CAG GTG

TAA AGA GGT CA-3′ (prim-177) and 5′-tac gaa ttc gag ctc AAG GAG CAG TTA GAT AAT GGA G-3′ (prim-179) and the DNA fragment for *pSN1301::MTD1-OsmiR156f* was obtained by primers prim-177 and 5′-ggg aaa ttc gag ctc AAG GAG CAG TTA GAT AAT GGA G-3′ (prim-178) and then infused the fragment of interest with *Kpn* I-*Sac* I of *pCAMBIA1301* and *pSN1301* by using ClonExpress II one-step cloning kit (Vazyme, C112-01) to generate plasmid *pCAM1301::MTD1-OsmiR156f* and *pSN1301::MTD1-OsmiR156f*, respectively. The mutant form sequences were obtained by an overlap extension polymerase chain reaction (PCR) method. To generate the *pCAM1301::△MTD1-OsmiR156f* in which *OsMTD1* gene sequence was deleted, primers 5′-gga tcc ccg ggt acc ctt aaa tgc tcc aat agc tag-3′ (prim-182) and prim-179 were used to amplify a fragment sequence from *OsMTD1* gene stop codon to the 60-bp sequence downstream of *pre-miR156f* from genomic DNA, and then the DNA fragment was ligated into the binary vector *pCAMBIA1301* for transformation. Similar strategies were carried out to construct *pCAM1301::ATT-OsmiR156f* in which the ATG start codon of *OsMTD1* was mutated to ATT. The primers 5′-gga tcc ccg ggt acc aga tcg ccg gag atT agc cag aag tc-3′ (prim-183) and prim-179 were used in the ATT mutant fragment amplification. For *pSN1301::△MTD1-OsmiR156f* and the *pSN1301::ATT-OsmiR156f*, a CaMV 35S promoter was harbored at the upstream of *pCAM1301::△MTD1-OsmiR156f* and *pCAM1301::ATT-OsmiR156f*, respectively. The corresponding primers prim-182 and prim-178 were employed for *pSN1301::△MTD1-OsmiR156f*, and prim-183 and prim-178 for *pSN1301::ATT-OsmiR156f*. The constructed vectors were infiltrated into the tobacco leaves by *Agrobacterium tumefaciens*–mediated transformation.

PCR, Real-Time PCR, and Stem-Loop RT-PCR

Genomic DNA was extracted and purified from fresh young leaves of five plants using CTAB methods. PCR was carried out in a reaction system with a total volume of 20 μL. The primers 5′-ATG AGC CAG AAG TCG TCG TGG C-3′ and 5′-ACA CAT GAA CGT ACA CGG CGC C-3′ were used for *OsMTD1* analysis. PCR validation for CNV was performed in all selected rice cultivars, and three independent experiments were performed for each cultivar. The primers were used as follows: primer64, 5′-AAA TGG CGG AAA CTT GAC AC-3′; primer65, 5′-TGA GCT AGC TGG ACA CAT GG-3′; primer66, 5′-CGG ACC TAA CCA CCG ATC TA-3′; primer67, 5′-ATC TTG GCG CTG CAA TTA TC-3′; inhF, 5′-ATG AGC CAG AAG TCG TCG TGG C-3′; inhR, 5′-ACA CAT GAA CGT ACA CGG CGC C-3′.

Total RNA was isolated from ∼100 mg leaves of five plants using a Trizol reagent (Invitrogen) and treated with RNase-free DNase I (Invitrogen) according to the manufacturer's instructions. Approximately 5 μg of RNA was used to synthesize first-strand cDNA using poly (dT) oligo primer according to the manufacturer's instructions in M-MLV kit (Invitrogen). Quantitative real-time reverse transcription PCR (RT-qPCR) was carried out in a reaction system with a total volume of 20 μL, which contained SYBR green I (Invitrogen) on a CFX96 system (BIO-RAD). The following programs were employed: predenaturing for 30 s at 95°C and then amplification for 40 cycles including denaturation for 10 s at 95°C, annealing for 30 s at 60°C, and extension at 72°C for 10 s. The *pri-miR156f* was normalized to the internal rice *tubulin*β*-4* gene, and the relative abundance was determined with $2^{-\Delta\ \Delta\ Ct}$ method. The RT-qPCR analysis in different lines was repeated three independent times. The primers for testing *pri-miR156f* were 5′-CTT CCC TTC GAC AGG ATA GC-30 and 5′-AGC GGC AGC TGT ATC ATC A-3′.

Stem-loop RT-qPCR (Varkonyi-Gasic et al., 2007) was employed to detect the mature osa-miR156f. Relative expression levels of osa-miR156 were normalized to the internal control *U6* in rice and *NbEF1* in tobacco. PCR was carried out in a reaction system with a total volume of 20 μL, which contained SYBR green I (Invitrogen) on a CFX96 system (BIO-RAD). The following programs were employed: predenaturing for 30 s at 95°C and then amplification for 40 cycles including denaturation for 10 s at 95°C, annealing for 30 s at 60°C, and extension at 72°C for 10 s. The $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression level of osa-miR156, and the analysis was repeated three independent times. The primers for *U6* are 5′-TAC AGA TAA GAT TAG CAT GGC CCC-3′ and 5′-GGA CCA TTT CTC GAT TTG TAC GTG-3′, and primers for *NbEF1* are 5′-GAT TGG TGG TAT TGG TAC TGT C-3′ and 5′-AGC TTC GTG GTG CAT CTC-3′.

RESULTS

OsMTD1-Located Segment Involves a New CNV in Rice

OsMTD1 sequence was queried *via* BLAST against four databases: TIGR rice genome annotation[1], Rice Information GateWay (RIGW[2]), National Center for Biotechnology Information[3], and the Knowledge-Based Oryza Molecular Biological Encyclopedia (KOME[4]). In the reference genome sequence of *O. sativa* spp. *japonica* cv. *Nipponbare*, we found that another locus, *Loc_Os08g34249*, is identical in DNA sequence to *OsMTD1*, a gene previously reported responsible for tillering in rice. However, it is not the case in the genomes of *indica* cultivars such as 93–11 (**Figure 1A**) and Shuhui498. Further analysis showed that *OsMTD1* and *Loc_Os08g34249* genes located on rice chromosome 8 according to their positions given in the TIGR rice database. *OsMTD1* and *Loc_Os08g34249* genes coexisted in two overlapped PAC clone AP0082414 and clone AP004703, implying one segmental duplication event on *OsMTD1*. To detect the physical location whereby the CNV event began and ended, we mapped the DNA sequences with different lengths between *OsMTD1* and *Loc_Os08g34249* against the reference genome and found the Nipponbare harbored a 13,002-bp tandem segmental duplication on *OsMTD1*-located region on the

[1]http://rice.plantbiology.msu.edu/

[2]http://rice.hzau.edu.cn/

[3]http://www.ncbi.nlm.nih.gov/

[4]http://www.cdna01.dna.affrc.go.jp/cDNA

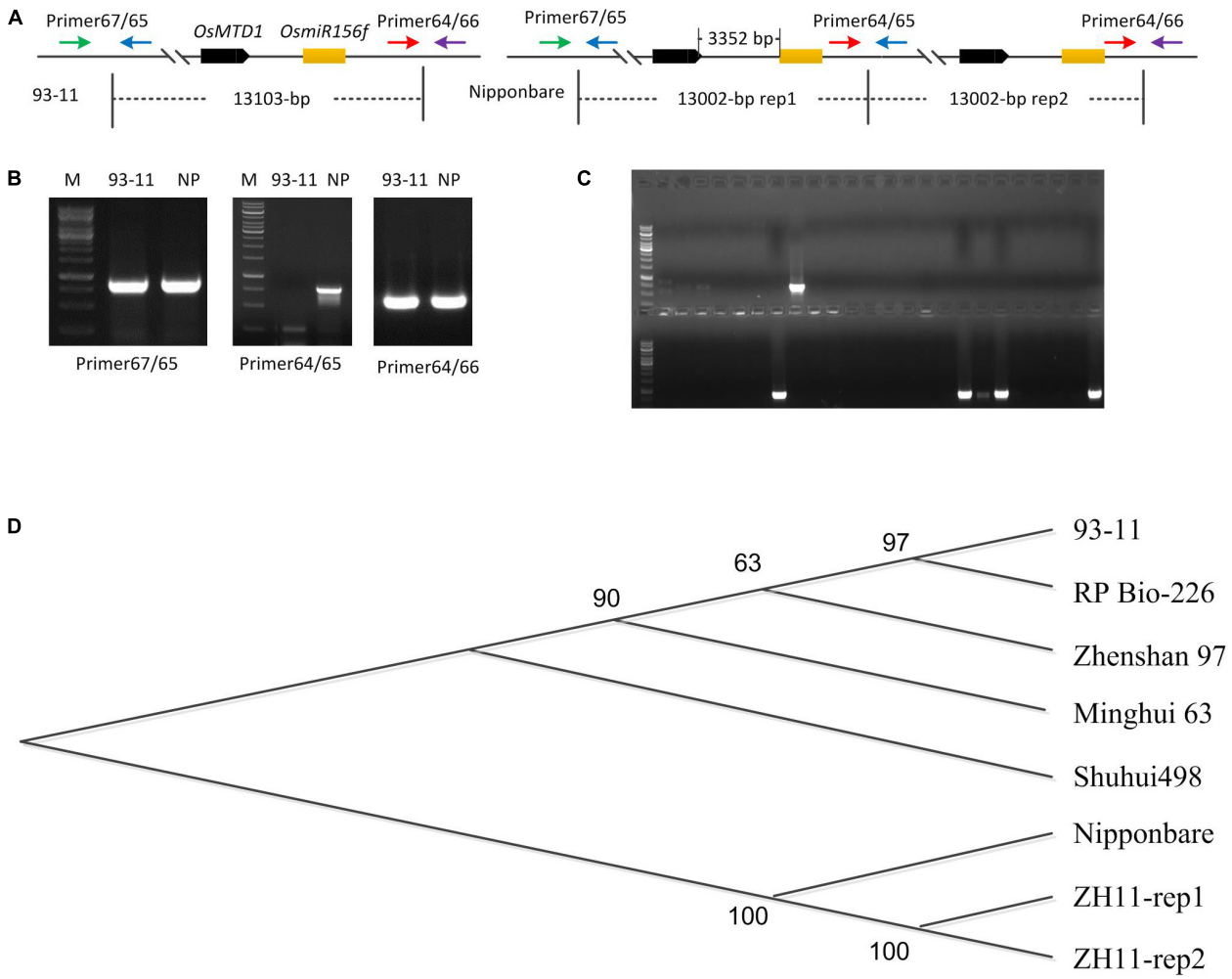

FIGURE 1 | *OsMTD1*-located segment involves a CNV. **(A)** The schematic shows the fragment of *OsMTD1* located on chromosome 8 in 93–11 and Nipponbare. Primer67/65, Primer64/66, and Primer64/65: primer pairs using for tandem duplication event analysis; *OsMTD1* and *OsmiR156f*: the two comprising elements in the corresponding DNA sequence of the *OsMTD1* located CNV. **(B)** Validity analysis on the primers detecting polymorphisms between 93–11 and Nipponbare. **(C)** Partial results from PCR amplification using different rice cultivars. **(D)** Phylogenetic analysis of CNV corresponding region's sequences in different rice cultivars. Nipponbare: the first sequence of the two DNA segments in *japonica* cv. Nipponbare genome; ZH11-rep1, ZH11-rep2: the first and the second sequence of the two DNA segments in *japonica* cv. ZH11 genome; Shuhui498-rep, 93–11-rep, Minghui 63, RP Bio-226, and Zhanshan97-rep: the corresponding DNA sequence of the *OsMTD1* located-CNV in different *indica* cultivars' genome.

eighth chromosome. *OsMTD1* and *Loc_Os08g34249* genes were reciprocal duplication, and each of them was encompassed in a 13,002-bp segment, respectively. Compared with the reference genome of Nipponbare, the tandem duplication in *indica* cultivar 93–11 is absent, and the varied length is more than 1 kb, so the 13,002-bp region encompassing *OsMTD1* could be regarded as a CNV between different rice cultivars. Herein, this DNA segment (about 13,002-bp corresponding region) variation in different rice cultivars was designated as *OsMTD1*-located CNV.

It was reported that the genome sizes of both *indica* and *japonica* subspecies have increased by greater than 2 and 6%, respectively, since their divergence from a common ancestor (Ma and Bennetzen, 2004). To find out whether this CNV contributes to intraspecific genome variations, PCR-amplified corresponding region was employed for a panel of 230 rice cultivars comprising both *indica* and *japonica* subspecies (**Supplementary Table 1**). The primers were designed according to the genomic sequence of both *japonica* cv. Nipponbare and *indica* cv. 93–11 to distinguish whether a tandem segmental duplication is harbored in the *OsMTD1*-located nearby region. A 754-bp fragment could be amplified from Nipponbare DNA with primer64 and primer65 but not from 93–11 (**Figure 1A**). The results showed that a clear band was obtained by two primer pairs (primer65 and primer67, primer64, and primer66) in all rice cultivars, representing the flanking sequences of the start or the end points of the corresponding 13,002-bp segment region in Nipponbare, respectively. However, a band was amplified with primer pair of primer64 and primer65 only in the ones whose genome harboring a tandem segmental duplication at the *OsMTD1* gene locus nearby region (**Figure 1B**). After validation by PCR, only

13 *japonica* cultivars including Nipponbare were found to have a tandem duplication in the corresponding region of *OsMTD1*-located segment (**Figure 1C** and **Supplementary Table 1**).

The tandem duplication of *OsMTD1*-located CNV corresponding sequence only appears in some *japonica* cultivars, but not in all investigated *indica* cultivars (**Figure 1C** and **Supplementary Table 1**); thus, this CNV represents a large inserted region only in some *japonica* cultivars. We then used the corresponding sequence of *OsMTD1*-located CNV from Nipponbare as a query to search against rice database for other cultivars deposited in National Center for Biotechnology Information (see text footnote 3), including *japonica* cv. Zhonghua11 (ZH11) and *indica* cv. Zhenshan 97, Minghui 63, 93–11, Shuhui 498, and RP Bio-226. Comparative analysis showed that the sequences of *OsMTD1*-located CNV region in different rice cultivars were highly conserved, and the dramatic divergences were found between *japonica* and *indica* subspecies (**Figure 1D** and **Supplementary Table 2**). The corresponding fragment of *OsMTD1*-located CNV region includes 107 SNPs, 10 deletions, and 11 insertions, resulting in 111-bp increase in *indica* cv. Shuhui 498, compared with Nipponbare. However, those regions are highly conserved in *indica* cultivars; it reaches 99.96% identity with only a 10-bp deletion and three SNPs among all five *indica* cultivars. As in Nipponbare, a tandem segmental replication at the *OsMTD1*-located regions is found in *japonica* cv. Zhonghua11. However, different from the complete sequence identity of two replication regions in Nipponbare, the sequences of the two DNA segments (designed as rep1 and rep2 according to the order occurred in genome) harbor 35 SNPs or mutations in ZH11, and the identities with the sequence of Nipponbare in rep1 and rep2 are 99.74 and 99.92%, respectively. Distance and Homology matrix analysis using the sequences of *OsMTD1*-located CNV further showed distant evolutionary relationships among different cultivars (**Supplementary Table 3**).

Phenotypic Difference According to *OsMTD1*-Located CNV

Copy number variations locating regions that contain protein-coding genes or important regulatory elements often have phenotypic effects (Żmien'ko et al., 2014). Our previous report showed that a T-DNA insertion in *OsMTD1* caused a dramatic change in tiller number (Liu et al., 2015). We therefore postulated this *OsMTD1*-located CNV has effects on a particular architecture trait, i.e., the tiller number. To investigate whether *OsMTD1*-located CNV affects tillering in rice, we performed phenotypic studies using 190 cultivars, including 108 *indica* and 82 *japonica*. Generally, *indica* and *japonica* cultivars show different tillering abilities, so the comparison between the two subspecies provided a reference to judge whether tiller number is a reliable trait for *OsMTD1*-located CNV conveying phenotype analysis. All cultivars were classified into four categories according to the tiller number: scale 1 (<10), scale 2 (10–20), scale 3 (21–30), and scale 4 (>30). Four *indica* while no *japonica* cultivars were classified into scale 4 (**Figure 2A**). On the contrary, more *japonica* cultivars were classified into scale 1 than *indica* cultivars, amounting to 55.5 and 22.5% of the investigated (**Figures 2A,C,D**), respectively. The average tiller number in *indica* was apparently higher than that in *japonica* cultivars (**Figure 2B**), indicating the selected 190 cultivars are a feasible representative group for tillering ability analysis. Further comparative analysis showed that the one-copy normal cultivars produced significantly increased tillers than the tandem duplicated cultivars (**Figure 2B**). In *japonica* cultivars, 6 of 55 one-copy normal cultivars (10.9%) showed tiller number of scale 3, whereas none was found in 13 two-copy cultivars (**Figures 2C,D**).

OsMTD1-Located CNV Involves in Rice Plant Architecture

The *OsMTD1*-located CNV region covers about 13,000 bp in different rice cultivars (**Supplementary Table 2**). GO analysis revealed that, apart from *OsMTD1*, an miR156 family member osa-miR156f was also contained in this CNV, and the *pre-miR156f* sequence located on downstream 3,352-bp away from *OsMTD1* (**Figure 1A**). The comparative analysis indicated the sequences of *OsMTD1* are identical in all cultivars, and *pre-miR156f* has identical sequences in all *indica* cultivars (Zhenshan 97, Minghui 63, 93–11, Shuhui 498, and RP Bio-226) but shows sequence differences in *japonica* cultivars (Nipponbare and ZH11); however, the final functional sequences of osa-miR156f and osa-miR156f* are completely identical in all investigated cultivars (**Supplementary Figure 1**). It is noteworthy that the osa-miR156 was confirmed to be positively correlated with rice tillering (**Supplementary Figure 2**; Schwab et al., 2005; Xie et al., 2006; Wang L. et al., 2015; Liu et al., 2019), whereas the tiller number comparison in cultivars with different CNV copies showed that the *OsMTD1*-located CNV region negatively affects rice tillering (**Figures 2B,D**). The osa-miR156f is aggressively antagonistic to *OsMTD1*-located CNV effect on tillering ability, suggesting that *OsMTD1* plays a vital role in the CNV.

To better understand the role of *OsMTD1* in the CNV, we further analyzed whether *OsMTD1* is directly involved in tiller development; CRISPR/Cas9 genome-editing technology was employed to generate both *Loc_Os08g34249* and *OsMTD1* knockout lines under ZH11 background. In 34 independent T_0 transgenic lines, sequence analysis revealed that each one belongs to heterogeneity accompanying an A/T/G/C insertion or deletion in *OsMTD1* (**Supplementary Figure 3**). Surprisingly, no double-knockout mutant was obtained in the CRISPR/Cas9 editing line after self-crossing for four times, and all CRISPR/Cas9 editing lines displayed no obvious phenotypic change. We further carried out CRISPR/Cas9 and overexpression analysis in one copy *japonica* cv. Kitaake and obtained many independent single-base deletion or insertion transgenic lines. All of the single-base mutation lines, in which an A/T/G/C was inserted or deleted, resulting in a frame-shift mutation and the original stop codon of *OsMTD1*, were excluded (**Supplementary Table 4**). Some mutation lines (such as line A-8 and A-44) significantly increased tiller number, whereas others (such as

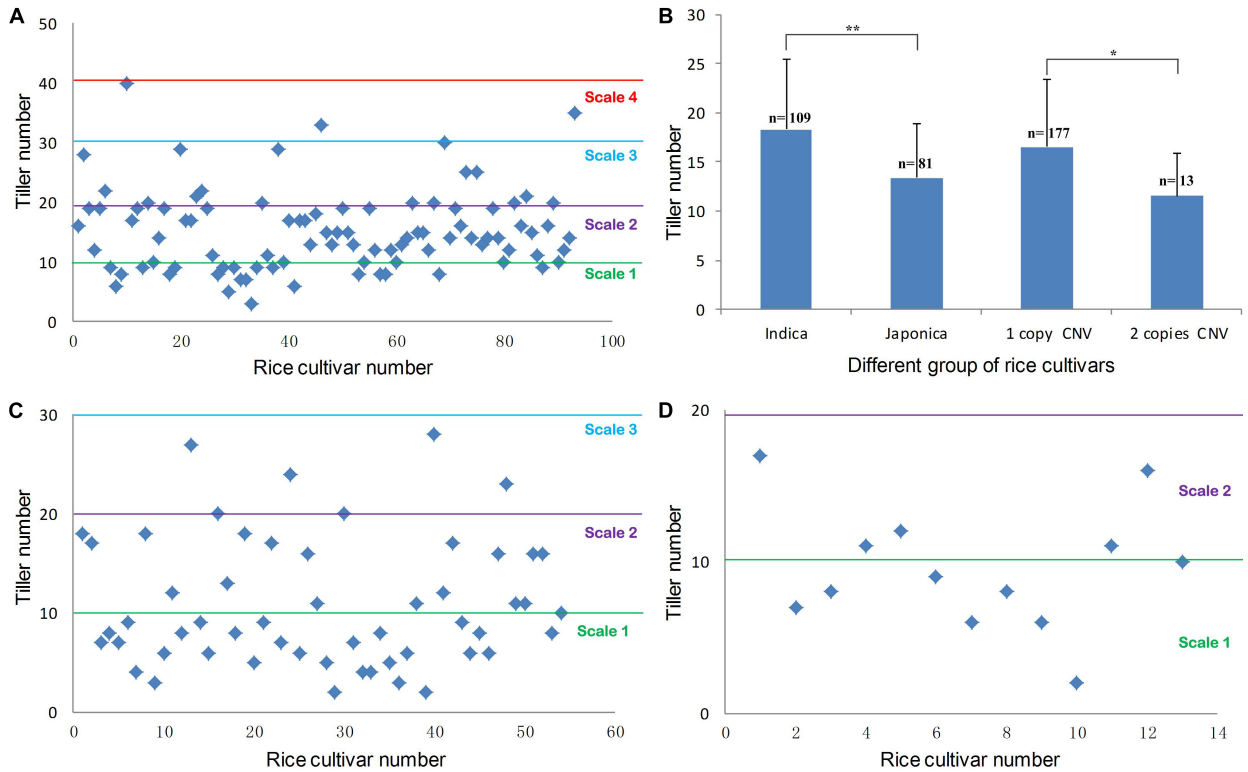

FIGURE 2 | Phenotypic distribution of rice tiller number. **(A)** The maximum tiller number from 94 *indica* accessions. **(B)** The statistical result of the maximum tiller number between *indica* and *japonica* accessions, the maximum tiller number between one-copy and two-copy CNV cultivars. Statistical significance was estimated by Student *t* tests. $^{*}P < 0.05$ and $^{**}P < 0.01$. **(C)** The maximum tiller number in different *japonica* accessions with one copy of *OsMTD1* nearby genome sequence. **(D)** The maximum tiller number in different *japonica* accessions with a tandem replication at *OsMTD1* nearby genome sequence.

A-3) showed no difference in tillering ability compared to the wild type. In fact, *OsMTD1* overexpression significantly decreased tiller number (**Figure 3**), even though the lines showed different *OsMTD1* increased levels (**Supplementary Table 5**). Intriguingly, the *OsMTD1* overexpression caused multiple phenotypic defects, such as reduction in grain number and plant height, whereas the height of *OsMTD1* CRISPR/Cas9 editing lines was comparable to that of the wild type plants (**Supplementary Figure 4**). These results under one-copy *OsMTD1*-located CNV background indicated that *OsMTD1* plays a prominent role in the genetic control of tillering ability in rice.

Leaf angle is an important agricultural trait determining rice plant architecture and ideotype (Zhou et al., 2017). In our experiments, the results suggest the pivotal role for *OsMTD1* in leaf inclination. Compared with the wild type, some *OsMTD1* CRISPR/Cas9 editing lines (such as A-8) showed no significant impact on leaf inclination at the mature stage, whereas overexpression of *OsMTD1* significantly reduced flag leaf angle (**Figure 4**). Consistently, the *OsMTD1* RNA*i* lines also presented increasing leaf inclination (**Supplementary Figure 2**). Conversely, overexpression of *pre-miR156f* increased leaf angle (**Supplementary Figure 5**). Meanwhile, leaf blades in miR156 knockout lines were found to be more erect than those of the wild type (Miao et al., 2019). These results also indicate that *OsMTD1* and osa-miR156 play opposite roles in regulating leaf angle. Taken together, it could be concluded that *OsMTD1*-located CNV contributes to a compact plant architecture by influencing both tiller number and leaf angle in rice.

OsMTD1 Inhibits the Transcript of *MicroRNA156f*

We next explored the underlying molecular mechanism of the *OsMTD1*-located CNV conveying phenotypes. Based on the experimental results mentioned above, an unexpected phenomenon is that the two elements or factors contained in the *OsMTD1*-located CNV region play opposite roles in controlling the architecture *via* tiller number and leaf angle: *OsMTD1* alone negatively regulates while osa-miR156f alone positively modulates these traits.

The short miRNAs (19–23 nt in length) are processed from corresponding large *pri-miRNAs*. In the large *pri-miRNAs* containing short open reading frame sequences that encode regulatory peptides, this miRNA-encoded peptide (miPEP) increasing the transcription of the *pri-miRNA* was reported (Lauressergues et al., 2015). Because of the close position of the *OsMTD1* to *pri-miR156f*, *OsMTD1* might regulate *pri-miR156f* transcription. To test this hypothesis, we first analyzed the osa-miR156 levels in both the CRISPR/Cas9 editing and

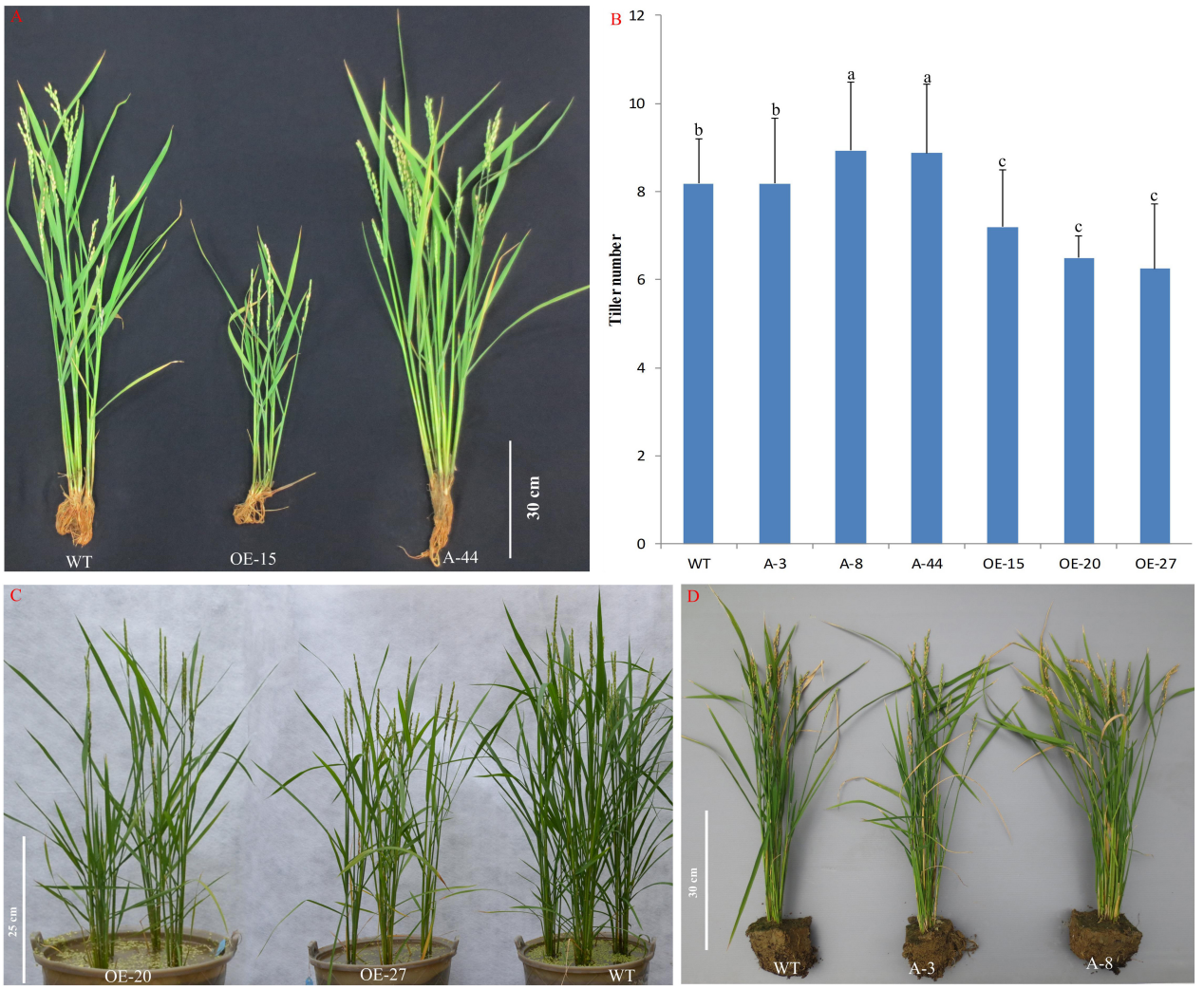

FIGURE 3 | Effects of *OsMTD1* on rice tillering. **(A)** Lines with different *OsMTD1* expression levels and the tillering ability performed. **(B)** Comparative analysis of the tiller number in lines with different *OsMTD1* expression levels. Statistical significance was estimated by Student *t* tests, and different letters indicate a significant difference ($P < 0.05$). **(C)** The *OsMTD1* overexpression lines produce less tillers compared to wild type. **(D)** The *OsMTD1* CRISPR/Cas9 editing lines show different tiller traits. WT: Kitaake; OE-15, OE-20, and OE-27: independent *OsMTD1* overexpression line; A-3, A-8, and A-44: Independent *OsMTD1* CRISPR/Cas9 editing line.

overexpression lines of *OsMTD1* under Kitaake background. The results showed that the CRISPR/Cas9 editing lines produced more osa-miR156 than the wild type (**Figure 5**). Among the *OsMTD1* overexpression plants, some lines (i.e., OE-15, OE-16, and OE-19) produced less osa-miR156, whereas some lines (i.e., OE-20, OE-26, and OE-27) produced comparable or more osa-miR156 than the wild type.

To further reveal regulatory role of OsMTD1 in osa-miR156f production, we used transformation of tobacco (*Nicotiana tabacum*) leaves to analyze the miR156f level by expressing both the native and mutant promoters of the *pre-miR156f* in *pri-miR156f*. Regardless of CaMV 35S promoter, compared with the amount of miR156f produced by expression of the native *pri-miR156f*, expression of an *OsMTD1* deletion mutant showed higher miR156 abundance. Likewise, expression of a *pri-miR56f* in which the ATG start codon of *OsMTD1* was mutated to ATT also produced higher miR156 level than expression of the native *pri-miR156f* (**Figure 5**), suggesting that OsMTD1 can inhibit the osa-miR156 accumulation when both are constructed in the same vector.

DISCUSSION

Copy number variations are major sources of genetic variation influencing gene expression and eventually the phenotype. It is believed that there are more CNVs than chromosome structural variations among individuals, and the total number of nucleotides covered by CNVs is much larger than SNP number in the whole genome (Lupski, 2007; Yu et al., 2013). CNVs can create new genes, change gene dosage, reshape gene structures, and modify elements regulating gene expression

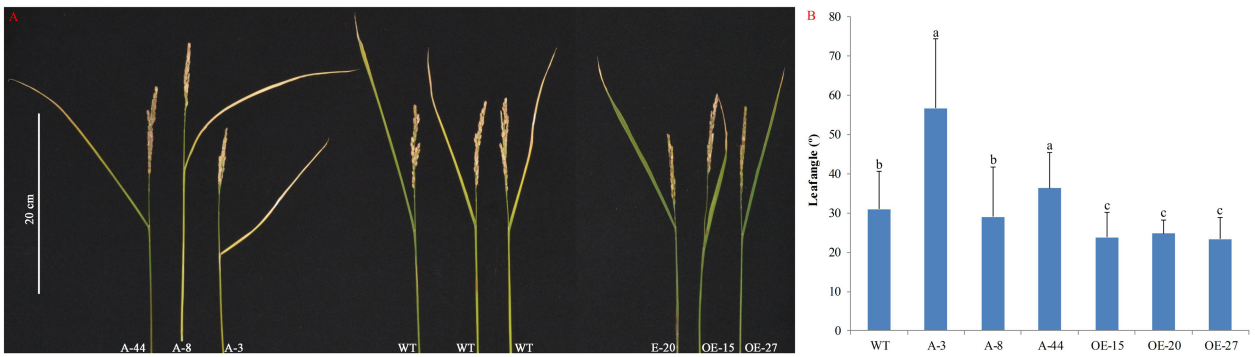

FIGURE 4 | Effects of *OsMTD1* on rice leaf angle. **(A)** Lines with different *OsMTD1* expression levels show different flag leaf angles. **(B)** Comparative analysis the flag leaf angle in lines with different *OsMTD1* expression levels. Statistical significance was estimated by Student *t* tests, and different letters indicate a significant difference ($P < 0.05$). WT: Kitaake; OE-15, OE-20, and OE-27: independent *OsMTD1* overexpression line; A-3, A-8, and A-44: independent *OsMTD1* CRISPR/Cas9 editing line.

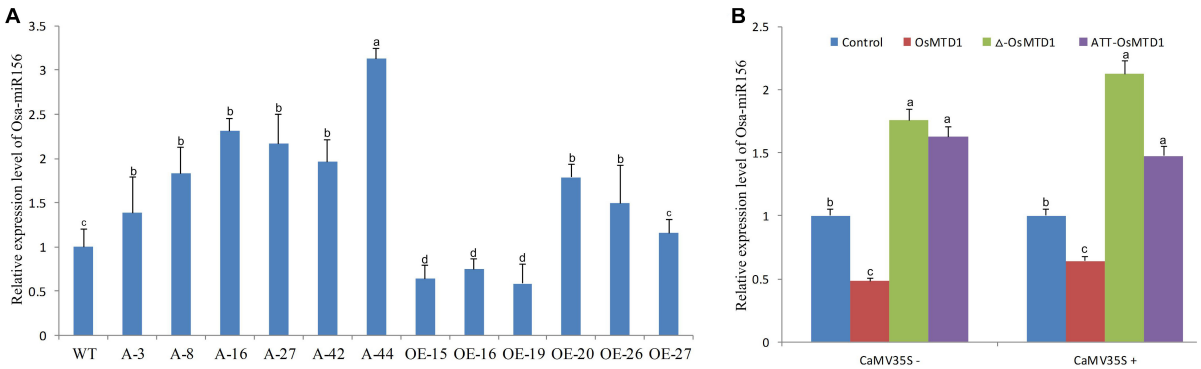

FIGURE 5 | Effects of *OsMTD1* on accumulation of miR156f. **(A)** The osa-miR156 relative levels in different rice lines. WT: Kitaake; A-3, A-8, A-16, A-27, A-42, and A-44: different *OsMTD1* CRISPR/Cas9 editing lines; OE-15, OE-16, OE-19, OE-20, OE-26, OE-27: different *OsMTD1* overexpression lines. **(B)** Quantification of miR156 in tobacco leaves expressing the *pri-miR156f* including *OsMTD1* sequence (OsMTD1), or the *pri-miR156f* in which the *OsMTD1* was deleted (Δ-OsMTD1), or in which the *OsMTD1* start codon was mutated to ATT (ATT-OsMTD1). The empty vector was used as control. Statistical significance was estimated by Student *t* tests, and different letters indicate a significant difference ($P < 0.05$).

(Henrichsen et al., 2009; Zhang et al., 2009). Here, we describe the identification of a new CNV, *OsMTD1*-located CNV, which involves an approximately 13,000-bp tandem duplication in DNA sequence on the eighth chromosome in different rice cultivars (**Supplementary Tables 1**, **2**), and the corresponding sequences of *OsMTD1*-located CNV region in different cultivars are highly conserved, and the sequence includes two important regulator factors, i.e., *OsMTD1* and *pri-miR156f*. The sequence of *OsMTD1*, osa-miR156f, and osa-miR156f* in all investigated cultivars is identified (**Supplementary Figure 1**).

Genome-scale studies indicated that CNVs significantly contribute to natural variation in plants (Yu et al., 2013; Żmien'ko et al., 2014; Bai et al., 2016). Changes in gene copy number provide the possibility to rapidly alter the dosage of a gene, which could directly cause a phenotypic variation, and as long as the new beneficial variation being selected over many generations under high selective pressure, the copy number alterations in a particular region may accumulate, and the phenotypic effects may intensify. Segmental duplications longer than 10 kb and of greater than ∼97% sequence identity can lead to local genomic instability (Stankiewicz and Lupski, 2010). As *OsMTD1*-located CNV covering approximately 13,000-bp DNA sequence is an evolutionarily recent duplication in some *japonica* cultivars after highly selective breeding programs, it is not surprising that *OsMTD1*-located CNV contributes to one or more currently advantageous traits in rice. In this article, we have investigated the tiller number on a panel of 190 rice cultivars, and results indicated that this CNV may have phenotypic effects on tiller development; for some, two-copy cultivars produced less tillers than one-copy cultivars (**Figure 3**). In the study, we also provided evidence that *OsMTD1*-located CNV contains two regulators, i.e., *OsMTD1* and osa-miR156f, jointly regulating tillering and leaf angle (**Figures 3**, **4**; **Supplementary Figures 2**, **4**, **5**). Together, these results indicate that *OsMTD1*-located CNV is important for rice plant architecture.

The essential role of a CNV depended on the genes or regulators contained in its region, so the roles of the *OsMTD1*-located CNV in rice phenotypes are determined by its two

comprising elements: *OsMTD1* and *pri-miR156f*. In order to reveal the function of *OsMTD1*, the first gene contained in *OsMTD1*-located CNV, CRISPR/Cas9-mediated gene editing technology was employed to knock out *OsMTD1* in the one-copy Kitaake and two-copy ZH11. Some lines both with multiple tillers and large leaf inclination were found under the Kitaake background (**Figures 3**, **4**). Although no obvious tillering phenotype change was observed in the CRISPR/Cas9 editing lines under the two-copy ZH11 background, the reason might be no knockout mutant was obtained in those experiments. Meanwhile, some *OsMTD1* overexpression lines that produced less tiller and smaller leaf angle were obtained (**Figures 3**, **4**). Combining the facts that both the *OsMTD1* RNA*i* lines and mutant lines with a T-DNA insertion into the region of this CNV in two-copy ZH11 displayed multitillering phenotypes (Liu et al., 2015), it is clear that *OsMTD1* alone could be regarded as an executive factor for tillering and leaf angle. For the second gene in *OsMTD1*-located CNV, i.e., *pri-miR156f*, we had demonstrated that the osa-miR156f plays crucial roles in rice tiller development (Liu et al., 2019), which is consistent with previous reports that high-level miR156 causes a bushy phenotype (Schwab et al., 2005; Xie et al., 2006). The effects of *OsmiR156f* on leaf angle are also verified (**Supplementary Figure 5**; Miao et al., 2019).

How *OsMTD1*-located CNVs regulate rice phenotype is another issue to be explored. It is believed that deletion and duplication can cause a phenotype change *via* several molecular mechanisms, and the commonly recognized mechanism is altering the copy number of a dosage-sensitive gene (or genes) (Lee and Lupski, 2006). The *OsMTD1*-located CNV enclosed two functional elements, *OsMTD1* and *pri-miR156f*; either can act alone as a pleiotropic regulator to determine rice plant architecture in a dosage-dependent manner. However, it seems further explanation is needed for the joint regulation mechanism in plant architecture by *OsMTD1*, *pri-miR156f*, and *OsMTD1*-located CNV. If the CNV phenotype is conveyed by altering *OsMTD1* and *pri-miR156f* dosage only, variation trends of the two contained genes should be the same – both increased or decreased along with the copy number change. In particular, transgenic experiments proved that changing two components of this CNV resulted in contradictory tillering phenotype: compared with the wild type, *OsMTD1* overexpression lines produced less tillers (**Figure 3**), whereas *pre-miR156f* overexpression lines significantly increased tillers (**Supplementary Figure 2**). Similarly, *OsMTD1* and *pre-miR156f* overexpressed lines also displayed opposite effects on leaf angle (**Figure 4**). All experimental data indicated that two genes contained in the CNV region, i.e., *OsMTD1* and *pri-miR156f*, play opposite roles alone in tiller number and leaf angle. Finally, the role of *OsMTD1*-located CNV in rice tillering and leaf angle is apparently consistent with *OsMTD1*, whereas contradictory to osa-miR156f, it was implied that the transcript of *OsMTD1* was more abundant in two-copy cultivars than in one-copy ones, whereas the opposite was true for *pri-miR156f*. Therefore, *OsMTD1* exhibits the major effect and acts as a key factor in the *OsMTD1*-located CNV region and thus contributes to a compact architecture in rice. One possibility is the different extent of genetic buffering, as *pri-miR156f* belongs to a large functionally redundant gene family, and the duplication in the *OsMTD1*-located CNV has only minor effects compared with *OsMTD1*. An alternative explanation is that there might be else unknown factors that inhibit *pri-miR156f* transcription in the CNV.

In addition to changes in gene dosage, many other mechanisms are responsible for the potential effects of CNVs, including reshaping of the gene structure and modification of the elements that regulate gene expression (Henrichsen et al., 2009; Zhang et al., 2009). One possible mechanism is the position effect; i.e., a CNV encompassed regulatory elements might regulate a gene even if they are several Mbs away (Żmien'ko et al., 2014). In the corresponding region of *OsMTD1*-located CNV, *OsMTD1* and *pri-miR156f* are neighboring genes approximately 3.3 kb apart (**Figure 1**). Hence, we reasoned that *OsMTD1* can inhibit the transcripts of *pri-miR156f* *via* position effect. Validation for the unpredictable effects of the two distant components in the CNV region is informative. We hypothesized that *OsMTD1* is a regulator repressing *pri-miR156f* transcription and provided some evidence. Compared with wild type, the *OsMTD1* CRISPR/Cas9 editing lines showed higher osa-miR156 level, whereas some *OsMTD1* overexpression lines showed lower osa-miR156 abundance (**Figure 5A**). Some *OsMTD1* overexpression lines didn't produce less osa-miR156 than wild type as expected, the reason might be that the insertion location is too far away from *pri-miR156f* in the genome. Furthermore, transformation results in tobacco leaves also showed that the native *pri-miR156f* vector produced less miR156 compared to the deleted and mutated types (**Figure 5B**). The above evidence implied that OsMTD1 can inhibit its neighboring *pri-miR156f* expression *in vivo* by the position effect. Different from previous report that miRNA-encoded peptide can enhance their corresponding *pri-miRNA* transcription (Lauressergues et al., 2015), OsMTD1 represses *pri-miR156f* transcription. Thus, our work revealed a novel regulatory mechanism for manipulating osa-miR156 level to control tiller number and leaf angle in rice.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

QL and LX conceived and designed the research. QL, JX, YZ, YM, and WK performed the experiments. QL, ZH, SX, YS, JT, CH, and WL analyzed the data. JX and X-FY investigated rice tiller number. C-ML provided technical support. QL, RW, KP, and LX wrote the manuscript. All authors read and approved the article.

FUNDING

This work was supported by the National Natural Science Foundation of China (91317312, 31900387, and 31570372), National Key Research and Development Program-Seven

Major Crops Breeding Project (2016YFD0101803), and Scientific Research Fund of Hunan Provincial Education Department (16K042).

ACKNOWLEDGMENTS

We thank Prof. Li-Jia Qu and Prof. Genji Qin (Peking University) for their valuable suggestions. We also thank Dr. Jin Miao, Dr. Dongshu Guo, Dr. Xiaoru Wei, Dr. Tong Wei, and Dr. Qing Tao (Peking University) for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.620282/full#supplementary-material

Supplementary Figure 1 | The sequence comparison result of *OsMTD1* and pre-*OsmiR156f* in different rice cultivars.

Supplementary Figure 2 | The CNV containing components *OsMTD1* and *OsmiR156f* influenced rice tiller development. WT: wild type; OsMTD1-RNAi: *OsMTD1* RNA interference line; OsmiR156f-OE: *OsmiR156f* overexpression line.

Supplementary Figure 3 | The *OsMTD1* mutagenesis analysis in rice using CRISPR/Cas9 system under ZH11 background.

Supplementary Figure 4 | Effects of *OsMTD1* on rice plant height. Statistical significance was estimated by Student *t* tests, and different letters indicate a significant difference ($P < 0.05$).

Supplementary Figure 5 | Effects of *OsmiR156f* on rice leaf angle.

Supplementary Table 1 | The surveying of CNV region and tiller number in different rice cultivars.

Supplementary Table 2 | DNA polymorphism analysis of *OsMTD1*-located CNV region in various rice cultivars. Nipponbare rep1, Nipponbare rep2: the first and the second sequence of the two DNA segments in Nipponbare genome; ZH11 rep1, ZH11 rep2: the first and the second sequence of the two DNA segments in ZH11 genome; Shuhui 498, 93–11, Minghui 63, RP Bio-226, and Zhenshan 97: the DNA sequence of the *OsMTD1*-located CNV in corresponding *indica* cultivar's genome.

Supplementary Table 3 | The result of Distance and Homology matrix analysis. Evolutionary analysis were conducted in DNAMAN8, and the analysis involved eight sequences: Nipponbare rep1: the first sequence of the two DNA segments in Nipponbare genome; ZH11-rep1, ZH11-rep2: the first and the second sequence of the two DNA segments in ZH11 genome; Shuhui 498, 93–11, Minghui 63, RP Bio-226, and Zhenshan 97: the DNA sequence of the *OsMTD1*-located CNV in corresponding *indica* cultivar's genome.

Supplementary Table 4 | The mutation information of *OsMTD1* in different CRISPR/Cas9 editing lines.

Supplementary Table 5 | The *OsMTD1* level in different overexpression lines.

REFERENCES

Bai, Z., Chen, J., Liao, Y., Wang, M., Liu, R., Ge, S., et al. (2016). The impact and origin of copy number variations in the *Oryza* species. *BMC Genom.* 17:261. doi: 10.1186/s12864-016-2589-2

Cong, B., Barrero, L. S., and Tanksley, S. D. (2008). Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. *Nat. Genet.* 40, 800–804. doi: 10.1038/ng.144

Feuk, L., Carson, A. R., and Scherer, S. W. (2006). Structural variation in the human genome. *Nat. Rev. Genet.* 7, 85–97.

Henrichsen, C. N., Chaignat, E., and Reymond, A. (2009). Copy number variants, diseases and gene expression. *Hum. Mol. Genet.* 18, R1–R8.

Hiei, Y., and Komari, T. (2008). Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nat. Protoc.* 3, 824–834. doi: 10.1038/nprot.2008.46

Hinds, D. A., Kloek, A. P., Jen, M., Chen, X., and Frazer, K. A. (2006). Common deletions and SNPs are in linkage disequilibrium in the human genome. *Nat. Genet.* 38, 82–85. doi: 10.1038/ng1695

Lauressergues, D., Couzigou, J.-M., Clemente, H. S., Martinez, Y., Dunand, C., Bécard, G., et al. (2015). Primary transcripts of microRNAs encode regulatory peptides. *Nature* 520, 90–93. doi: 10.1038/nature14346

Lee, J. A., and Lupski, J. R. (2006). Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron* 52, 103–121. doi: 10.1016/j.neuron.2006.09.027

Li, S., Wang, S., Deng, Q., Zheng, A., Zhu, J., Liu, H., et al. (2012). Identification of genome-wide variations among three elite restorer lines for hybrid-Rice. *PLoS One* 7:e30952. doi: 10.1371/journal.pone.0030952

Li, Y., Xiao, J., Wu, J., Duan, J., Liu, Y., Ye, X., et al. (2012). A tandem segmental duplication (TSD) in green revolution gene *Rht-D1b* region underlies plant height variation. *New Phytol.* 196, 281–291.

Liu, Q., Shen, G., Peng, K., Huang, Z., Tong, J., Kabir, M. H., et al. (2015). The alteration in the architecture of a T-DNA insertion rice mutant *osmtd1* is caused by up-regulation of *MicroRNA156f*. *J. Integrat. Plant Biol.* 57, 819–829. doi: 10.1111/jipb.12340

Liu, Q., Su, Y., Zhu, Y., Peng, K., Hong, B., Wang, R., et al. (2019). Manipulating *osa-MIR156f* expression by *D18* promoter to regulate plant architecture and yield traits both in seasonal and ratooning rice. *Biol. Proced.* 21:21.

Locke, D. P., Sharp, A. J., McCarroll, S. A., McGrath, S. D., Newman, T. L., Cheng, Z., et al. (2006). Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. *Am. J. Hum. Genet.* 79, 275–290. doi: 10.1086/505653

Lupski, J. R. (1998). Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.* 14, 417–422. doi: 10.1016/s0168-9525(98)01555-8

Lupski, J. R. (2007). Genomic rearrangements and sporadic disease. *Nat. Genet.* 39, S43–S47.

Ma, J., and Bennetzen, J. L. (2004). Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 12404–12410. doi: 10.1073/pnas.0403715101

Miao, C., Wang, Z., Zhang, L., Yao, J., Hua, K., Liu, X., et al. (2019). The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nat. Commun.* 10:3822.

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., et al. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.* 23, 1233–1236.

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev. Cell* 8, 517–527. doi: 10.1016/j.devcel.2005.01.018

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., et al. (2004). Large-scale copy number polymorphism in the human genome. *Science* 305, 525–528. doi: 10.1126/science.1098918

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40, 1023–1028. doi: 10.1038/ng.169

Stankiewicz, P., and Lupski, J. R. (2010). Structural variation in the human genome and its role in disease. *Annu. Rev. Med.* 61, 437–455. doi: 10.1146/annurev-med-100708-204735

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E. F., and Hellens, R. P. (2007). Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* 3:12. doi: 10.1186/1746-4811-3-12

Wang, L., Sun, S., Jiye, J., Fu, D., Yang, X., Weng, X., et al. (2015). Coordinated regulation of vegetative and reproductive branching in rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15504–15509. doi: 10.1073/pnas.1521949112

Wang, Y., Xiong, G., Hu, J., Jiang, L., Yu, H., Xu, J., et al. (2015). Copy number variation at the *GL7* locus contributes to grain size diversity in rice. *Nat. Genet.* 47, 944–948. doi: 10.1038/ng.3346

Weckselblatt, B., and Rudd, M. K. (2015). Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet.* 31, 587–599. doi: 10.1016/j.tig.2015.05.010

Xie, K., Wu, C., and Xiong, L. (2006). Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice. *Plant Physiol.* 142, 280–293. doi: 10.1104/pp.106.084475

Yu, P., Wang, C., Xu, Q., Feng, Y., Yuan, X., Yu, H., et al. (2013). Genome-wide copy number variations in *Oryza sativa* L. *BMC Genom.* 14:649. doi: 10.1186/1471-2164-14-649

Zhang, F., Gu, W., Hurles, M. E., and Lupski, J. R. (2009). Copy number variation in human health, disease, and evolution. *Annu. Rev. Genom. Hum. Genet.* 10, 451–481.

Zhang, L., Yu, H., Ma, B., Liu, G., Wang, J., Wang, J., et al. (2017). A natural tandem array alleviates epigenetic repression of *IPA1* and leads to superior yielding rice. *Nat. Commun.* 8:14789.

Zhou, L., Xiao, L., and Xue, H. (2017). Dynamic cytology and transcriptional regulation of rice lamina joint development. *Plant Physiol.* 174, 1728–1746. doi: 10.1104/pp.17.00413

Żmien'ko, A., Samelak, A., Kozłowski, P., and Figlerowicz, M. (2014). Copy number polymorphism in plant genomes. *Theor. Appl. Genet.* 127, 1–18. doi: 10.1007/s00122-013-2177-7

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 12 February 2021
doi: 10.3389/fpls.2021.640135

Histone Demethylases ELF6 and JMJ13 Antagonistically Regulate Self-Fertility in Arabidopsis

*Charlie Keyzor†, Benoit Mermaz†‡, Efstathios Trigazis, SoYoung Jo‡ and Jie Song**

Department of Life Sciences, Imperial College London, London, United Kingdom

OPEN ACCESS

Edited by:
Raju Datla,
Global Institute for Food Security (GIFS), Canada

Reviewed by:
Keqiang Wu,
National Taiwan University, Taiwan
Beth Allyn Krizek,
University of South Carolina, United States

***Correspondence:**
Jie Song
j.song@imperial.ac.uk

†These authors have contributed equally to this work

‡Present address:
Benoit Mermaz
Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences, Yale University, New Haven, CT, United States
SoYoung Jo
Center for Molecular Biomedicine, Friedrich-Schiller-Universität Jena, Jena, Germany

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

Received: *10 December 2020*
Accepted: *21 January 2021*
Published: *12 February 2021*

Citation:
Keyzor C, Mermaz B, Trigazis E, Jo S and Song J (2021) Histone Demethylases ELF6 and JMJ13 Antagonistically Regulate Self-Fertility in Arabidopsis. Front. Plant Sci. 12:640135. doi: 10.3389/fpls.2021.640135

The chromatin modification H3K27me3 is involved in almost every developmental stage in Arabidopsis. Much remains unknown about the dynamic regulation of this histone modification in flower development and control of self-fertility. Here we demonstrate that the H3K27me3-specific demethylases ELF6 and JMJ13 antagonistically regulate carpel and stamen growth and thus modulate self-fertility. Transcriptome and epigenome data are used to identify potential targets of ELF6 and JMJ13 responsible for these physiological functions. We find that ELF6 relieves expansin genes of epigenetic silencing to promote cell elongation in the carpel, enhancing carpel growth and therefore encouraging out-crossing. On the other hand, JMJ13 activates genes of the jasmonic acid regulatory network alongside the auxin responsive SAUR26, to inhibit carpel growth, enhance stamen growth, and overall promote self-pollination. Our evidence provides novel mechanisms of self-fertility regulation in *A. thaliana* demonstrating how chromatin modifying enzymes govern the equilibrium between flower self-pollination and out-crossing.

Keywords: chromatin regulation, histone modification, histone demethylases, epigenetics, flower development, self-fertility

INTRODUCTION

As a predominantly self-fertilizing plant, the growth and development of the male and female organs in the Arabidopsis flower need to be coordinated. How this is achieved is not yet fully understood (Wellmer et al., 2013). Since chromatin regulation is involved in almost every developmental process of a plant's life cycle, we investigated how it contributes toward the correct timing of floral organ development to enable self-fertility. A particular histone modification that is highly dynamic throughout Arabidopsis development is the trimethylation of the lysine 27 residue of histone 3 (H3K27me3) which induces transcriptional silencing (Francis et al., 2004; Entrevan et al., 2016; Frerichs et al., 2019). H3K27me3 is deposited at thousands of Arabidopsis genes by the polycomb repressive complex 2 (PRC2) (Zhang et al., 2007; Lafos et al., 2011) and represses floral development genes in the seedling (Wang et al., 2016). Accordingly, at some developmental stage these floral development genes must be reactivated by removal of H3K27me3 and addition of active chromatin marks such as H3K4me3 or H3K36me3 (Pfluger and Wagner, 2007).

Three genes have been demonstrated to encode targeted H3K27me3 specific demethylases which may reactivate these floral development genes; ELF6, REF6, and JMJ13 (Lu et al., 2011; Crevillén et al., 2014; Yan et al., 2018). Recent findings suggest that each demethylase is recruited to a large number of target genes, some of which are targeted by more than one demethylase (Yan et al., 2018; Antunez-Sanchez et al., 2020). Recruitment to these genes is achieved by a combination of

direct DNA binding via a Zinc finger domain and interaction with other transcription factors (Yu et al., 2008; Li et al., 2016; Yan et al., 2018). Certain physiological functions have already been assigned to the three demethylases, such as regulation of flowering time (Zheng et al., 2019), control of leaf cell elongation (Yu et al., 2008) and resetting the epigenome across generations (Crevillén et al., 2014; Antunez-Sanchez et al., 2020; Borg et al., 2020). Functions such as controlling leaf cell elongation require specific targeting of the demethylases to a subset of their global target genes. In the context of leaf cells, this is achieved by the interaction of ELF6 and REF6 with the BZR2 transcription factor which recruits the demethylases to specific target genes (Yu et al., 2008). Furthermore, in floral buds REF6 has been demonstrated to interact with a number of developmentally important MADS-box transcription factors (Yan et al., 2018).

The role of the demethylases in epigenetic reactivation of floral development genes and thus control of floral development is poorly understood. Though changes to floral morphology have been observed in *elf6 jmj13 ref6* triple mutants (Yan et al., 2018), the function of each individual demethylase in floral development and self-pollination control is unknown. In this study we reveal that two of the histone demethylases, ELF6 and JMJ13, antagonistically regulate self-pollination by modulating the growth of stamen and carpel, linking ELF6/JMJ13-dependent chromatin regulation to floral development and self-fertility. We further investigate the transcriptome and epigenome changes caused by loss of these demethylases to predict the target genes which may be responsible for these novel developmental functions.

MATERIALS AND METHODS

Plant Material

A. thaliana Columbia-0 ecotype (Col-0) was used in this study as wild-type material. All knock-out mutants were T-DNA insertions of Col-0: *jmj13* (GABI_113B06), *elf6-3* (SALK_074694C), *ref6-1* (SALK_001018C). Double and triple mutants were generated by crossing and genotyping (kindly provided by Prof C Dean, John Innes Centre, UK). pJMJ13::JMJ13-GFP in *jmj13* genetic background and p35S::ELF6-GFP in *elf6* genetic background (Kindly provided by Dr. H Yang and Prof C Dean, John Innes Centre, UK) were used for phenotypic complementation.

Fertility Assessment

The number of failed siliques was counted on the primary inflorescence after approximately 6 weeks of growth such that at least 10 siliques had matured on each primary inflorescence sampled.

Floral Organ Phenotype Measurements

Col-0, *elf6, jmj13, ref6,* all double and triple mutants were grown for 4–5 weeks until flowering. Stage 14 buds, as defined by Smyth et al. (1990), corresponded to the 2 youngest open buds of the inflorescence and were used as samples for floral organ height quantification. Buds were imaged using either the Leica MZ165 or the Olympus SZ61. ImageJ (Rueden et al., 2017) was used to perform measurements of stamens and carpels.

Pollen Viability Assay

Pollen from stage 14 flowers the first four flowers were stained with 5 ug/ml of fluorescein diacetate to examine pollen viability ($n = 5$ flowers per genotype, $n > 150$ pollen grains per genotype).

Transcriptome Sequencing

RNA was extracted from stage ∼9–13 buds (Smyth et al., 1990) by phenol-chloroform extraction (Box et al., 2011). Single-end deep sequencing of two replicates from each genotype was performed after mRNA enrichment (BGI technology). Sequences (>30 million from each sample) were aligned to the TAIR10 genome using bowtie2 (Langmead and Salzberg, 2012) and differentially expression analysis performed using NOISeq (Tarazona et al., 2011). Differentially expressed genes are defined by a 2-fold change in expression and a NOISeq probability score ≥ 0.8.

Ploidy Analysis

Ploidy analysis method was adapted from Yang et al. (2019). Sixteen carpels for each genotype were dissected from stage 14 flowers at inflorescence positions 1 and 2. The dissected carpels were immediately placed in ∼400 ul of nuclei isolation buffer (0.01 M MgSO4, 0.05 M KCl, 1.2 mg/ml HEPES buffer, 10 mM DTT, 2.5% Triton-X100 in water). Carpel tissue was then diced in the buffer to release the nuclei and was filtered through a double layer of Mira-cloth (pore size 22–25 μm). DAPI was added to a final concentration of 2 μg/ml and incubated at room temperature for 15 min. This solution (400 μl) was then run through the BD LSR Fortessa™ flow cytometer and DAPI was excited with a 405 and 640 nm laser and light collected at 450 nm+/−25 and 780 nm+/−30, respectively.

Cell Elongation Analysis

Flowers from Col-0 ($n = 4$), *elf6* ($n = 3$), and *jmj13* ($n = 2$) were sampled at stage 14 from the first flower position on the inflorescence. Each flower was dissected to leave just the intact gynoecium attached to the stem. The gynoecium was cleared in 80% isopropanol for between 110 and 130 min and then stained in 20 μg/ml propidium iodide (PI) solution for 20 min. The PI stained gynoeciums were imaged with a Leica SP5 upright confocal microscope with an excitation wavelength of 514nm and an emission capture range of 585–602 nm. Cross-sections of the upper and lower carpel were imaged. The cells of the outer epidermal cell files of the carpel were measured along the long axis of the carpel in ImageJ (Rueden et al., 2017), measuring at least 95 cells per genotype.

General Data Analysis

The R statistical programming language (R Core Team, 2019) was used for all data analysis and graph generation excluding the flow cytometry data which was analyzed in FCSalyzer (https://sourceforge.net/projects/fcsalyzer/). The following R packages were utilized in the analysis and graphing of data: tidyverse (Wickham et al., 2019), eulerr (Larsson, 2020), BioMaRt (Durinck et al., 2009).

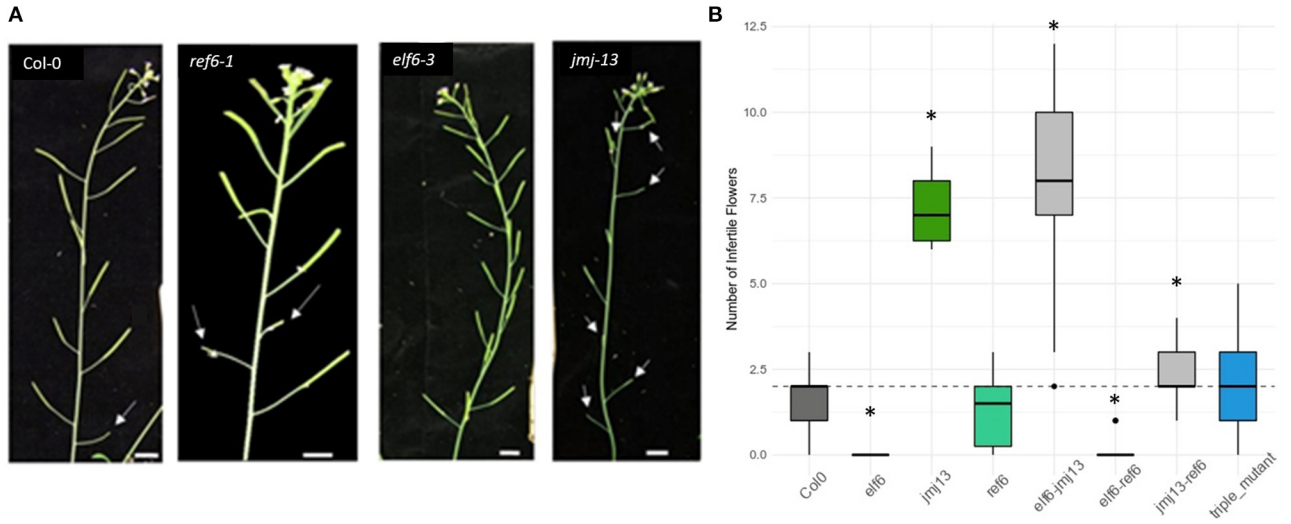

FIGURE 1 | Fertility phenotypes of H3K27me3 demethylase knock-out mutants. **(A)** Varying degrees of self-fertility in the demethylase knock-out mutants. The white arrow heads point to aborted siliques caused by failed self-pollination, as seen in the first flower of Col-0 and to the eighth flower of *jmj13*. Scale bars represent 1 cm. **(B)** Quantification of self-fertility. Numbers of aborted siliques are plotted with asterisks indicating which data point is statistically significant as compared to that of Col-0 ($p < 0.05$, *t*-test). The dashed horizontal line represents the median *y*-axis value of Col-0.

Gene Accession Numbers

A. thaliana gene locus identification codes for genes mentioned in this study (for the full list of predicted target genes see **Supplementary Table 1**) are as following: ELF6 (AT5G04240); REF6 (AT3G48430); JMJ13 (AT5G46910); BZR1 (AT1G75080); BZR2 (AT1G19350); INO (AT1G23420); MYB24 (AT5G40350); SMR8 (AT1G10690); EXPA1 (AT1G69530); EXPA3 (AT2G37640); EXPB3 (AT4G28250); SAUR26 (AT3G03850); AGP14 (AT5G56540); AGP22 (AT5G53250); AGP7 (AT5G65390).

Data Availability

All RNA sequencing datasets generated in this study can be accessed through NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/gds) under accession number GSE164739.

RESULTS

ELF6 and JMJ13 Antagonistically Regulate Arabidopsis Self-Fertility via Floral Organ Growth

Fertility in the context of Arabidopsis refers to the ability of a flower to generate a full-length silique with viable seeds. Arabidopsis thaliana is predominantly self-fertilizing, in that the female gynoecium of a typical flower is fertilized by pollen from the same flower. However, the first two flowers to mature often fail to self-fertilize resulting in stunted siliques (**Figure 1A**). The infertility in these first two flowers is caused by a failure to self-pollinate, as the stigma extends beyond the reach of the mature stamens preventing pollen transfer from anther to stigma (**Figure 2A**). By manually fertilizing flowers with their own pollen, we restored fertility confirming that the absence of self-fertility is not caused by gamete viability but by failure to transfer pollen (**Supplementary Figure 1**).

To reveal the role of the H3K27me3 demethylases in regulating self-pollination, the number of infertile siliques was quantified in *elf6, ref6,* and *jmj13* T-DNA mutants (**Figure 1**) (see methods for specific alleles). In *elf6* and *jmj13*, significant and opposite changes to the degree of self-fertilization were observed. *elf6* displayed increased self-fertility whereby all flowers, even the first two flowers, were consistently self-fertile. Conversely, an infertility phenotype was observed in *jmj13* whereby aborted siliques were observed all the way to the eighth flower of the primary inflorescence. ref6 did not show a significant change in fertility. The relationship between the demethylases was also probed by quantifying fertility in double and triple mutant combinations. Double mutants displayed non-obvious floral phenotypes; *elf6 jmj13* adopted a *jmj13*-like reduced fertility, *elf6 ref6* was super-fertile while *jmj13 ref6* showed an intermediate phenotype between *jmj13* and *ref6*. Finally, the triple mutant showed no significant change from the wild-type (**Figure 1B**, **Supplementary Figure 2**).

To confirm that the fertility phenotypes are attributed to loss of the demethylases rather than artifacts of random T-DNA insertions, stable complementation lines expressing p35S::ELF6-GFP and pJMJ13::JMJ13-GFP in *elf6* and *jmj13* mutant backgrounds, respectively, were characterized. Both transgenic lines restored fertility to wild-type levels (**Supplementary Figure 3**). We next questioned whether the infertility phenotype of *jmj13* may be caused by a reduction in pollen viability. Pollen from Col-0, *elf6* and *jmj13* was stained with fluorescein diacetate in which only viable pollen displays fluorescein fluorescence. No significant

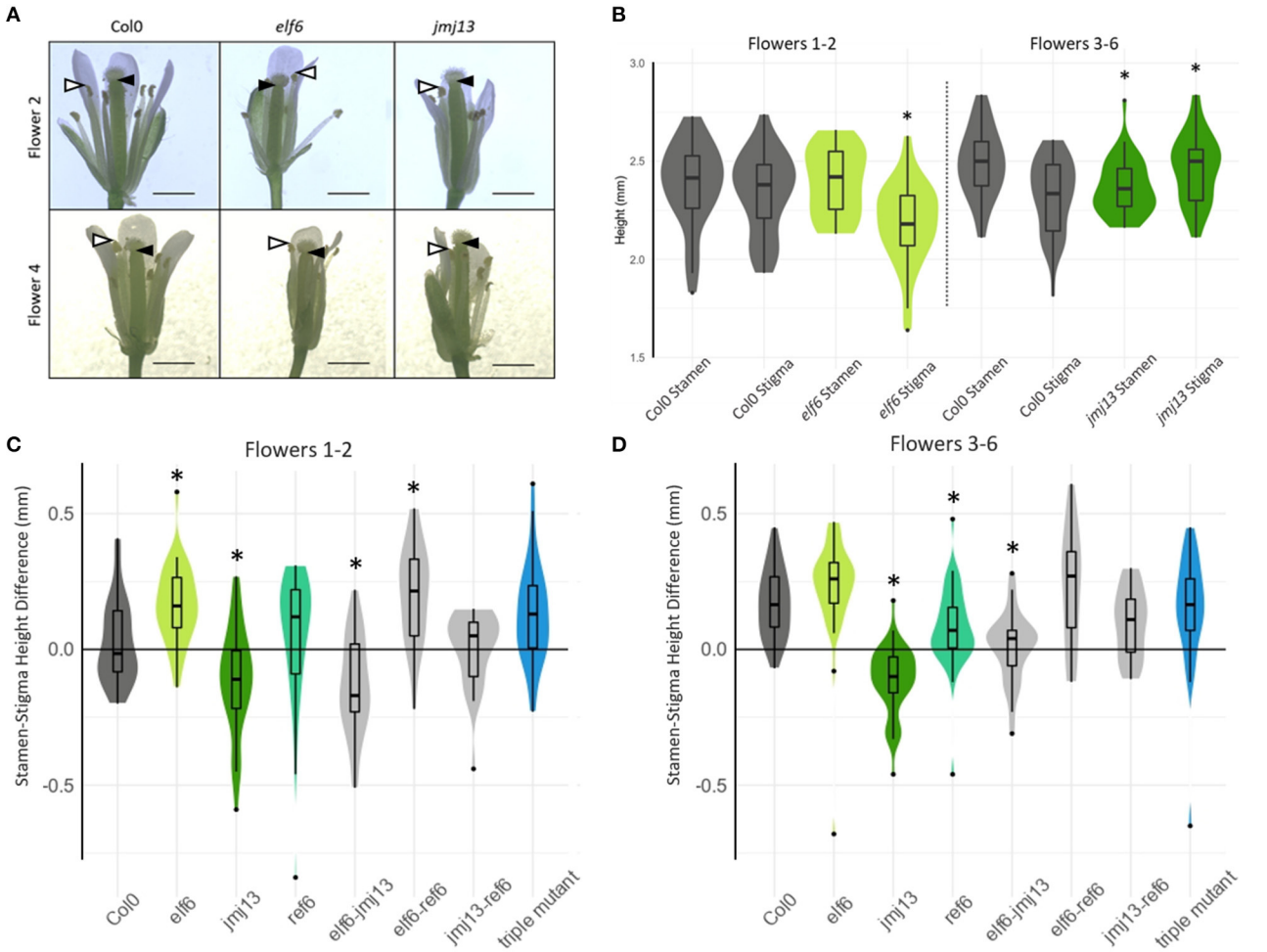

FIGURE 2 | Quantification of floral organ lengths in H3K27me3 knock-out mutants. **(A)** Representative images of stage 14 flowers of the indicated genotypes from flower positions 2 and 4 (scale bars represent 1 mm). Black arrow heads indicate the lowest point of the stigma closest to the receptacle, whilst white arrow heads indicate highest point of the tallest stamen, the difference of which correlates with self-fertility. **(B)** The absolute heights of stamen and stigma from Col-0, *elf6,* and *jmj13* flowers. Measurements were taken from receptacle to the top of the highest stamen and to the bottom of the stigma. $n >= 20$ for each genotype in each flower position. An asterisk indicates statistical significance compared to the respective Col-0 data of the same floral organ and flower position group ($p <= 0.05$, *t*-test). **(C)** The lengths from receptacle to the tip of the tallest stamen, and receptacle to the lowest point of the stigma were measured. The difference between these two lengths is plotted for every flower measured at flower positions 1–2 and distributed by genotype. Groups marked with an asterisk are statistically significantly different from the Col-0 stamen/stigma height differences ($p < 0.05$, *t*-test). The violin graph represents the density of data point distribution displayed symmetrically along the y axis. $n >= 20$ for each genotype. **(D)** Stamen-stigma height differences were measured and plotted as in **(C)** but for flower positions 3–6. $n >= 20$ for each genotype.

change in pollen viability was found between Col-0 and *jmj13* or *elf6* (**Supplementary Figure 4**). To verify that male and female reproductive organs were still individually functional in the absence of JMJ13, manual self-pollination was performed. Fertility of the first two flowers was restored in Col-0 and *jmj13* when they were manually self-pollinated (**Supplementary Figure 1**, $n \geq 36$).

We hypothesized that these changes in fertility upon loss of H3K27me3 demethylases were due to changes in carpel and stamen growth altering the probability of self-pollination. To assess this hypothesis, the lengths of carpel and stamen were measured from flowers at developmental stage 14 when self-pollination typically occurs in Col-0 (Smyth et al., 1990) (**Figure 2A**). Two lengths were measured: from the receptacle (base of the flower) to the top of the tallest stamen and from the receptacle to the bottom of the stigma (**Figure 2B**). The difference between these two lengths was calculated such that positive values indicate that the stamens were taller than the stigma and self-pollination was likely to occur, whereas negative values indicate the stamens were shorter than the stigma and the flower was likely to fail in self-pollination. As expected, Col-0 displayed a negative median floral organ height difference in the first two flowers and positive median difference in flowers 3–6 (**Figure 2A**) as is consistent with the typical fertility of flowers in those positions (**Figure 1B**). The stamen-stigma height differences in the single mutants were also consistent with their

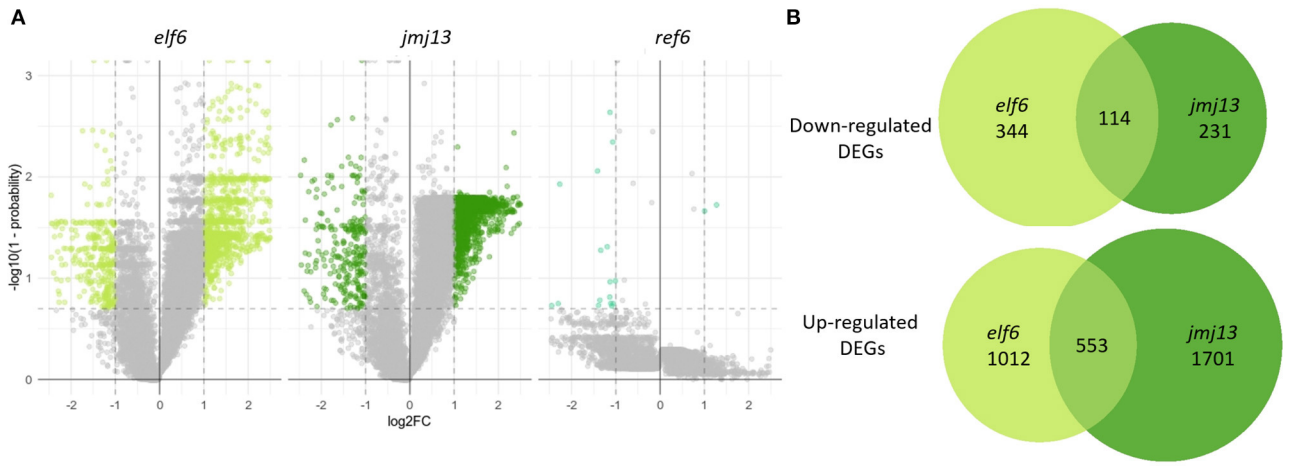

FIGURE 3 | Transcriptome changes in the bud in *elf6, ref6* and *jmj13* knock-out mutants compared to Col-0. **(A)** Transcriptomes of *elf6, jmj13,* and *ref6* buds compared after differential expression analysis against the transcriptome of Col-0 buds. The significance thresholds drawn in dashed lines represent a 2-fold change in expression and a probability score > 0.8. Genes over these thresholds are classed as differentially expressed and colored in lime green, forest green and cyan dependent on the genotype. **(B)** DEGs extracted from the *elf6* and *jmj13* transcriptomes. They are split into up and down regulated DEGs and represented as individual venn groups, with the shared group representing genes that are either differentially up or down regulated in both *elf6* and *jmj13* buds compared to Col-0.

fertility phenotypes. The "super-fertile" *elf6* displayed a positive median stamen/stigma height difference in all flower positions and was significantly greater than Col-0 in flowers 1–2 ($p < 0.05$, *t*-test); the semi sterile *jmj13* median stamen/stigma difference was negative and significantly less than Col-0 in all flower positions measured ($p < 0.05$, *t*-test). *ref6* median stamen-stigma difference showed no statistically significant difference from Col-0 in flowers 1–2. Specifically, the absolute heights of stamen and stigma show that *elf6* had significantly shorter carpels than Col-0 in flowers 1–2 whilst *jmj13* had significantly longer carpels and shorter stamens than Col-0 in flowers 3–6 (**Figures 2B,C**). The data demonstrates that JMJ13 and ELF6 play crucial antagonistic roles in regulating self-fertility, by regulating the growth of specific floral organs.

Transcriptome Wide Changes Occur in *elf6* and *jmj13* Mutants

To understand the role of H3K27me3 demethylation in regulating floral organ growth and self-pollination, transcriptome datasets were generated from stage ∼9–13 buds of Col-0, *elf6*, *jmj13,* and *ref6* inflorescences. Large transcriptome changes were observed in *elf6* and *jmj13*; defined by > 2-fold expression change and > 0.8 NoiSeq probability score, there were 2,023 and 2,599 differentially expressed genes (DEGs), respectively, when compared to Col-0 (**Figure 3**). The *ref6* knock-out buds on the other hand displayed only minor transcriptome changes as only 103 DEGs were observed, consistent with the subtle fertility phenotype of *ref6*. There is proteomic evidence to show that REF6 binds to a number of important MADS-box transcription factors controlling floral development (Yan et al., 2018) and so it is unexpected that *ref6* displays the weakest floral development phenotype and transcriptome effects. One possible explanation may be that the T-DNA insertion in the *ref6-1* mutant (SALK_001018C) does not completely inhibit REF6 function. This may be because in this *ref6-1* mutant the catalytic domain remains intact and only the zinc-finger domain is disrupted by the inserted T-DNA. Previous studies (Yan et al., 2018) have shown that REF6 still targets several thousand genes without its zinc-finger domain, merely losing specificity, and so *ref6-1* likely also retains targeting and catalytic ability. Read alignment to the *ref6* gene model confirmed that the catalytic domain is still expressed in the *ref6* T-DNA insertion mutant (**Supplementary Figure 5**). This is in contrast to the *elf6* and *jmj13* mutants in which the catalytic domain is completely disrupted by T-DNA insertion.

The *elf6* and *jmj13* transcriptome datasets were further analyzed to gain a mechanistic understanding of self-fertility regulation by the H3K27me3 demethylases. DEGs from both the *elf6* and *jmj13* datasets contained mostly up-regulated genes. This demonstrates that the majority of DEGs are not direct floral targets of the demethylases as a floral target gene would remain epigenetically silenced without the respective H3K27me3 demethylase due to ectopic accumulation of H3K27me3. Significant, but minority, overlap was found between the *elf6* and *jmj13* DEGs (3230 unique DEGs, 696 shared DEGs, $p < 0.0001$).

A data screening approach was taken to predict ELF6 and JMJ13 target genes responsible for the floral organ growth phenotypes. We first assumed that without ELF6 or JMJ13, their respective target genes would become ectopically enriched in H3K27me3. Using a H3K27me3 ChIP-seq dataset from Col-0 and demethylase triple mutant buds (Yan et al., 2018) we selected all genes which showed significant H3K27me3 enrichment in the triple mutant buds (3,216 genes). The phenotypic target genes would also be expected to be transcriptionally silenced in the *elf6* or *jmj13* mutant due to H3K27me3 accumulation. From our RNAseq data, 61 and 42 genes from the 3,216 gene subset were significantly down-regulated in *elf6* and *jmj13,* respectively. This gene list comprises our predicted ELF6 and

TABLE 1 | Predicted ELF6 phenotypic target genes as determined by screening of transcriptome and epigenome data.

ELF6 target gene name	Target gene function	Log2(*elf6*/Col-0)	Log2(*jmj13*/Col-0)
EXPA1	Promotes cell elongation by cell wall loosening (Cosgrove 2015)	−1.26	0.47
EXPA3	Promotes cell elongation by cell wall loosening (Cosgrove 2015)	−1.51	−0.49
EXPB3	Promotes cell elongation by cell wall loosening (Cosgrove 2015)	−1.01	0
ADP1/ABS3	Promotes cell elongation, branching, sensecence and fertility in auxin dependent manner (Li et al., 2014; Wang et al., 2015)	−1.1	−0.16
PRR9	Regulation of the circadian clock and flowering time (McClung and Gutiérrez, 2010)	−1.22	1.02
ZFP2	Negatively regulates floral abscission and stamen length (Cai and Lashbrook, 2008)	−3.19	−1.2
CBF1	Transcriptional activator of cold tolerance and ABA response genes (Li et al., 2017)	−2.76	0.38
CYP81F4	Secondary metabolite synthesis	−1.66	0.16
MARD1	Regulator of ABA mediated seed dormancy and interactor of SnRK1 autophagy activator (He and Gan, 2004; Nietzsche et al., 2014)	−1.39	0.63
LEA2	Promotes root growth and drought tolerance (Magwanga et al., 2018)	−1.8	−0.52
EPFL2	Stomata guard cell differentiation	−1.07	0.67
FLZ7	Adaptor for SnRK1 autophagy regulation (Jamsheer et al., 2018)	−2.12	−1.01
ATL98	Ubiquitin mediated protein degradation	−2.3	0.09
MBOAT	Membrane bound O-acyl transferase	−1.05	0.19
PLC6	Phosphoinositol-DAG signaling	−1.1	0.26
AT2G29660	Zinc-finger transcription factor	−1.07	0.12
CUAOy2	Amine metabolism	−1.3	0.47
DTX49	Xenobiotic efflux	−1.5	−0.24

JMJ13 target genes. To further reduce this list to the most likely candidates it was assumed that because *elf6* and *jmj13* show opposing fertility and carpel growth phenotypes that the genes responsible should be differentially expressed between the *elf6* and *jmj13* mutants. 41 of the 103 predicted ELF6/JMJ13 target genes showed a 2-fold change in expression between *elf6* and *jmj13* (**Supplementary Table 1**). From this subset, genes of unknown function were discarded and the remaining 18 ELF6 target genes (**Table 1**) and 11 JMJ13 target genes (**Table 2**) were determined to be the most likely ELF6/JMJ13 target genes causing the observed phenotypes.

ELF6 Likely Relieves Expansin Genes of H3K27me3 Suppression to Induce Cell Elongation

Of the 18 predicted phenotypic target genes of ELF6, three were identified as expansins (EXPB3, EXPA3, and EXPA1) (**Supplementary Figure 8**). The expansins are cell wall remodeling enzymes which disrupt non-covalent bonding between cellulose microfibrils to relax the cell wall allowing turgor-pressure to induce cell-elongation (Cosgrove, 2015). Expansins induce cell elongation responses in numerous developmental contexts (Marowa et al., 2016), but have yet to be implicated in carpel growth. It is likely that the loss of ELF6 has caused multiple expansin genes to remain epigenetically silenced by H3K27me3 and thus unable to induce cell elongation in the carpel, retarding carpel growth. We next asked how ELF6 might be recruited to the expansin genes. The expansins are shown to be regulated by the brassinosteroid response transcription factor BZR2 which has also been demonstrated to directly bind ELF6 and REF6 (Yu et al., 2008, 2011). It is likely that in the carpel BZR2 recruits ELF6 to multiple expansin genes. Supporting this theory, we found that our 60 predicted ELF6 target genes are significantly enriched in ChIP-seq validated target genes of BZR2 and its close homolog BZR1 whereas the predicted JMJ13 target genes showed no such enrichment (ELF6-BZR2, x2.26 enrichment, $p = 0.049$; ELF6-BZR1, x1.78 enrichment, $p = 0.026$, hypergeometric test, **Supplementary Figure 6**).

If expansin epigenetic silencing in *elf6* is the primary cause of stunted carpel growth, we should observe a reduction in cell size in the *elf6* carpel. Cell elongation is a key growth mechanism for the maturing silique (Ripoll et al., 2019). However, it is unclear whether differential cell elongation at the self-pollination stage can affect the chances of successful pollen deposition onto the stigma. Using confocal microscopy, we measured cell length in the carpel epidermis of stage 14, position 1 flowers from Col-0, *elf6* and *jmj13* plants (**Figure 4**). A significant 15% decrease in mean cell length was observed in *elf6* carpels compared to Col-0 carpels (n carpels ≥ 3, n cells ≥ 95, $p <= 0.05$, *t*-test), from a mean cell length of 24.5 μm in Col-0 (SD = 7.6) to 20.8 μm in *elf6* (SD = 6.9). This reduction in cell length could explain the ∼8% reduction in total carpel length of position 1 flowers (Col-0 mean length = 2.51 mm, SD = 0.15; *elf6* mean length = 2.32 mm, SD = 0.10) (**Figure 2B**), therefore, supporting a role for the expansins in promoting carpel growth to prevent self-pollination.

TABLE 2 | Predicted JMJ13 phenotypic target genes as determined by screening of transcriptome and epigenome data.

JMJ13 target gene name	Target gene function	Log2(*elf6*/Col-0)	Log2(*jmj13*/Col-0)
JAZ7	Regulator of jasmonic acid signaling, flower development, flowering time, drought tolerance, pathogen defense (Browse and Wallis, 2019)	1.43	−1.77
SAUR26	Auxin responsive positive regulator of cell elongation, enriched expressed in stamens (Spartz et al., 2014)	−0.19	−2.13
AGP22	Membrane proteoglycan involved in cell-cell signaling and wound response (Guan and Nothnagel, 2004)	−1.1	−2.29
AGP14	Membrane proteoglycan involved in cell-cell signaling and wound response (Guan and Nothnagel, 2004)	0.25	−1.17
AGP7	Membrane proteoglycan involved in cell-cell signaling and wound response (Guan and Nothnagel, 2004)	0.64	−1.2
LEC	Lectin up-regulated by chitin, mechanical damage, jasmonic acid, ethylene	2.24	−2.24
SAQR	Regulator of flowering time and starch allocation (Jones et al., 2016)	−0.1	−3
XTH24	Cell wall expansion by xyloglucan cleavage and re-ligation (Lee et al., 2018)	0.46	−1.37
PUP18	Purine and possibly cytokinin transporter, induced by the AP3 and PI homeotic transcription factors (Mara and Irish, 2008)	1.8	−1.94
ATL89	Ubiquitin mediated protein degradation	−0.14	−1.94
UPS4	Ureide permease	−0.31	−2.45

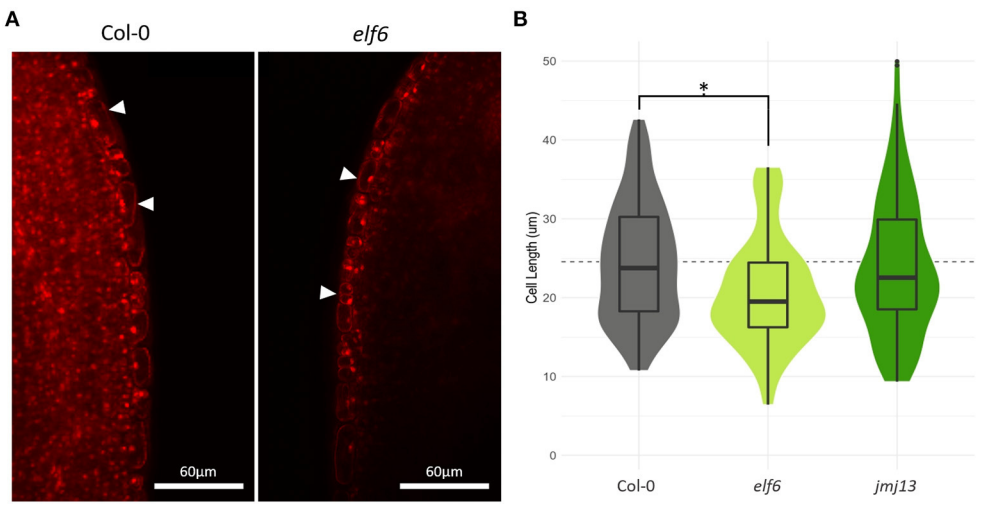

FIGURE 4 | Effects of *elf6* knock-out on carpel epidermal cell length. **(A)** Confocal microscopy of stage 14 carpel mid sections using propidium iodide staining to highlight individual cells. The image on the left is taken from a Col-0 carpel, whilst the right hand image is from *elf6.* White arrow heads indicate the epidermal cell layer. Scale bar depicts 60 µm. **(B)** Measurements of cell length in carpel epidermis. A significant 15% reduction in mean epidermal cell length was observed in *elf6* carpels compared to Col-0 carpels ($n >= 3$ carpels per genotype, $n >= 95$ cells measured per genotype, $p <= 0.05$, *t*-test). The dashed line represents the mean carpel epidermal cell length of Col-0 carpels. The asterisk denotes a significant change in cell length between the indicated genotypes ($p <= 0.05$, *t*-test).

Our bioinformatic pipeline also proposed the cyclin dependent kinase inhibitor SMR8, a likely regulator of endoreduplication, to be targeted by ELF6 (Van Leene et al., 2010) (log2FC(*elf6*/*jmj13*) = −0.9, so not featured in **Table 1**). We hypothesized that endoreduplication, mediated by SMR8, may be a synergistic growth mechanism in the wildtype carpel and that loss of ELF6 could silence SMR8 thus inhibiting endoreduplication and carpel growth. To verify this hypothesis, the ploidy state of stage 14 gynoecium nuclei was assessed using DAPI staining and flow cytometry. No significant endoreduplication could be detected in either Col-0, *elf6,* or *jmj13* gynoecia, concluding that endoreduplication is unlikely to be employed in carpel growth up to floral developmental stage 14 and is not affected by ELF6 or JMJ13 in the carpel (**Supplementary Figure 7**).

JAZ7, SAUR26 and Multiple Arabinogalactan Proteins Are Silenced in jmj13 Buds

Of particular interest in the set of predicted JMJ13 phenotypic target genes is the jasmonic acid response transcription factor JAZ7 (**Supplementary Figure 8**). Jasmonic acid is a key hormonal regulator of floral organ development and acts by binding JAZ repressor proteins to induce their degradation (Huang et al., 2017). JAZ proteins bind and inhibit other developmental transcription factors (Song et al., 2011) such as YABBY and MYB transcription factors to bring about controlled floral development (Meister et al., 2005; Reeves et al., 2012; Boter et al., 2015; Qi et al., 2015; Gross et al., 2018). Two members of the YABBY family in particular, CRC and INO, are involved in gynoecia development and both display stunted carpel growth in knock-out mutants (Meister et al., 2005). INO was found to be significantly up-regulated in the *jmj13* inflorescence (1.4 log2FC, 0.98 probability score) which may be considered phenotypically consistent with the enlarged gynoecia observed in *jmj13* mutants (**Figure 2B**). However, it is not clear whether this dysregulation of INO expression is related to the ectopic epigenetic silencing of JAZ7.

The other common JAZ targets, the MYB transcription factors, are also heavily involved in floral development. MYB21 and MYB24 in particular are known to promote stamen elongation and *myb21-myb24* knock-out mutants display a reduced stamen/gynoecium length ratio and significantly decreased fertility in a very similar manner to the *jmj13* knock-out phenotype (Reeves et al., 2012; Qi et al., 2015). MYB24 was found to be significantly down-regulated in the *jmj13* knock-out inflorescence (−1.6 log2FC, 0.99 probability score), providing a phenotypically consistent hypothesis to explain the reduced stamen growth and subsequent infertility of *jmj13* knock-out flowers. However, there were no significant changes to the H3K27me3 profile across the MYB24 gene, indicating that MYB24 must be indirectly repressed downstream of a JMJ13 target. Several JAZ proteins (JAZ1, JAZ8, JAZ11) have been shown to bind the MYB family transcription factor MYB24 (Song et al., 2011) and so it may be possibly be a target of JAZ7 too. Though there is lacking evidence to tie together the observed ectopic H3K27me3 silencing of JAZ7 and the gene expression changes of MYB24 and INO, it is clear that JMJ13 is playing an important role within this regulatory network.

In addition to the jasmonic acid regulator JAZ7, the auxin response gene SAUR26 is also epigenetically silenced in *jmj13* (**Supplementary Figure 8**). Numerous SAUR genes have been demonstrated to induce cell elongation via cell wall acidification (Spartz et al., 2014) and the SAUR63 subfamily has been specifically demonstrated to induce stamen elongation (Chae et al., 2012). Though the role of SAUR26 in stamen elongation has not been studied, SAUR26 does show highly enriched expression in the stamens and so is likely to serve the same function as the SAUR63 subfamily (Klepikova et al., 2016).

A final major class of potential JMJ13 target genes is the arabinogalactan proteins (AGPs); highly glycosylated proteins of the outer plasma membrane implicated in developmental signaling (Seifert and Roberts, 2007). Three AGPs were predicted to be phenotypic targets of JMJ13; AGP7, AGP14, and AGP22. However, the biochemical function and physiological role of specific AGPs is poorly understood and difficult to relate to floral organ growth. Interestingly, perturbation of AGPs elicits a wound-like response (Guan and Nothnagel, 2004) hinting at a possible connection between JMJ13, AGPs and the jasmonic acid regulatory network which also regulates wound response.

DISCUSSION

Stamen/carpel growth coordination is differentially regulated according to the position of flowers along the length of the inflorescence, such that a small portion of flowers refrain from self-fertilization (Plackett et al., 2017). Despite failing to self-pollinate, the stamen and gynoecium of these first flowers are still fully functional and able to produce viable siliques if pollen is artificially transferred from anther to stigma (**Supplementary Figure 1**). It is likely that *A. thaliana* in the wild receives some form of evolutionary advantage by refraining from self-pollination in the first flowers. In the wild, pollinators are likely to cause pollen transfer and thus enable out-crossing in these first two flowers providing evolutionary benefits in the form of genetic diversity. However, self-pollination is a far safer reproduction strategy as it is less dependent on external pollinators and resources are not invested by the plant to attract pollinators as with other species. Therefore, by self-pollinating all but the first two flowers, *A. thaliana* likely optimizes the benefits of out-crossing and self-fertilization (Stebbins, 1974; Wright et al., 2013). Our evidence demonstrates that the histone demethylases ELF6 and JMJ13 antagonistically regulate this evolutionary equilibrium between self-fertilization and out-crossing. We find that ELF6 reactivates floral development genes to promote carpel growth and outcrossing, whilst JMJ13 reactivates a different set of floral development genes to inhibit stamen growth and promote carpel growth to stimulate self-pollination.

Transcriptome and epigenome data (Yan et al., 2018) from flowers of demethylase mutants has been used to specifically identify the likely target genes causing these floral organ growth effects. Our findings suggest that ELF6 epigenetically activates multiple expansin genes, via BZR2 recruitment, inducing carpel cell elongation. The identification of JAZ7 as a JMJ13 target gene implicates chromatin regulation as a mechanism mediating the jasmonic acid gene regulatory network. The current literature is unclear on what the downstream effects of JAZ7 epigenetic silencing might be, largely due to there being multiple possibly redundant JAZ genes hindering studies on single gene knock-out mutants (Wager and Browse, 2012). Our observation of MYB24 down-regulation provides a phenotypically consistent explanation as to why *jmj13* displays stunted stamens and decrease fertility, but MYB24 does not show H3K27me3 enrichment in the demethylase triple mutant implying that it is not a direct demethylase target gene. Similarly with INO; although INO and its binding partner CRC are known to play a role

in promoting carpel growth (Gross et al., 2018), INO is clearly not a direct target of JMJ13 as it is up-regulated in the *jmj13* mutant. As JAZ7, MYB24, and INO are part of the same regulatory network, it is possible that an uncharacterised intermediate connects JAZ7 function to transcriptional regulation of MYB24 and INO. Moreover, the complex feedback of the floral development regulatory network suggests that ectopic epigenetic silencing of a single gene such as JAZ7 may induce unexpected changes to gene expression throughout the network (Reeves et al., 2012). An alternative or synergistic hypothesis linking JMJ13 to stamen growth is the epigenetic activation of SAUR26, a gene likely to induce stamen elongation via cell wall acidification (Chae et al., 2012).

We have demonstrated that the histone demethylases ELF6 and JMJ13 epigenetically regulate distinct sets of floral development genes to regulate floral morphology. Our evidence supports mechanisms whereby ELF6 promotes carpel elongation via epigenetic activation of expansin genes, whilst JMJ13 represses carpel growth via jasmonic acid signaling and promotes stamen growth via epigenetic reactivation of SAUR26. These conclusions establish histone demethylation as a key mechanism in regulating the chromatin state of floral development genes and hence controlling floral morphology and the equilibrium between self-pollination and out-crossing.

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: NCBI Gene Expression Omnibus, accession no: GSE164739".

AUTHOR CONTRIBUTIONS

CK, BM, and JS contributed to the conception and design of the study. CK, BM, ET, SJ, and JS performed the experiments and analyzed the data. CK performed all bioinformatics analysis and drafted the manuscript. CK, BM, and JS revised and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Royal Society under awards UF130276, RG140567, and RGF\EA\180140 and the Biotechnology and Biological Sciences Research Council under award BB/M011178/1.

ACKNOWLEDGMENTS

We would like to thank Prof Caroline Dean and Dr. H Yang (John Innes Centre, UK) for kindly providing double and triple demethylase mutant lines and ELF6-GFP and JMJ13-GFP transgenic lines.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.640135/full#supplementary-material

REFERENCES

Antunez-Sanchez, J., Naish, M., Ramirez-Prado, J. S., Ohno, S., Huang, Y., Dawson, A., et al. (2020). A new role for histone demethylases in the maintenance of plant genome integrity. *Elife* 9:e58533. doi: 10.7554/eLife.58533.sa2

Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendía, D., Axelsson, E., et al. (2020). Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat. Cell Biol.* 22, 621–629. doi: 10.1038/s41556-020-0515-y

Boter, M., Golz, J. F., Giménez-Ibañez, S., Fernandez-Barbero, G., Franco-Zorrilla, J. M., and Roberto Solano, R. (2015). Filamentous flower is a direct target of JAZ3 and modulates responses to jasmonate. *Plant Cell* 27, 3160–3174. doi: 10.1105/tpc.15.00220

Box, M. S., Coustham, V., Dean, C., and Mylne, J. S. (2011). Protocol: A simple phenol-based method for 96-well extraction of high quality RNA from Arabidopsis. *Plant Methods* 7:7. doi: 10.1186/1746-4811-7-7

Browse, J., and Wallis, J. (2019). Arabidopsis flowers unlocked the mechanism of jasmonate signaling. *Plants* 8:285. doi: 10.3390/plants8080285

Cai, S., and Lashbrook, C. (2008). Stamen abscission zone transcriptome profiling reveals new candidates for abscission control: enhanced retention of floral organs in transgenic plants overexpressing arabidopsis ZINC FINGER PROTEIN2. *Plant Physiol.* 146, 1305–1321. doi: 10.1104/pp.107.110908

Chae, K., Isaacs, C., Reeves, P., Maloney, G., Muday, G., Nagpal, P., et al. (2012). Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation. *Plant J.* 71, 684–697. doi: 10.1111/j.1365-313X.2012.05024.x

Cosgrove, D. (2015). Plant expansins: diversity and interactions with plant cell walls. *Curr. Opin. Plant Biol.* 25, 162–172. doi: 10.1016/j.pbi.2015.05.014

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., et al. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature* 515:587–590. doi: 10.1038/nature13722

Durinck, S., Spellman, P., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* 4, 1184–1191. doi: 10.1038/nprot.2009.97

Entrevan, M., Schuettengruber, B., and Cavalli, G. (2016). Regulation of genome architecture and function by polycomb proteins. *Trends Cell Biol.* 26, 511–251. doi: 10.1016/j.tcb.2016.04.009

Francis, N. J., Kingston, R. E., and Woodcock, C. L. (2004). Chromatin compaction by a polycomb group protein complex. *Science* 306, 1574–1577. doi: 10.1126/science.1100576

Frerichs, A., Engelhorn, J., Altmüller, J., Gutierrez-Marcos, J., and Werr, W. (2019). Specific chromatin changes mark lateral organ founder cells in the Arabidopsis inflorescence meristem. *J. Exp. Bot.* 70, 3867–3879. doi: 10.1093/jxb/erz181

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

Guan, Y., and Nothnagel, E. (2004). Binding of arabinogalactan proteins by yariv phenylglycoside triggers wound-like responses in Arabidopsis cell cultures. *Plant Physiol.* 135, 1346–1366. doi: 10.1104/pp.104.039370

He, Y., and Gan, S. (2004). A novel zinc-finger protein with a proline-rich domain mediates ABA-regulated seed dormancy in Arabidopsis. *Plant Mol. Biol.* 54, 1–9. doi: 10.1023/B:PLAN.0000028730.10834.e3

Huang, H., Liu, B., Liu, L., and Song, S. (2017). Jasmonate action in plant growth and development. *J. Exp. Bot.* 68, 1349–1359. doi: 10.1093/jxb/erw495

Jamsheer, K. M., Sharma, M., Singh, D., Mannully, C., Jindal, S., et al. (2018). FCS-like zinc finger 6 and 10 repress SnRK1 signalling in Arabidopsis. *Plant J.* 94, 232–245. doi: 10.1111/tpj.13854

Jones, D., Zheng, W., Huang, S., Du, C., Zhao, X., Yennamalli, R., et al. (2016). A clade-specific Arabidopsis gene connects primary metabolism and senescence. *Front. Plant Sci.* 7:983. doi: 10.3389/fpls.2016.00983

Klepikova, A., Kasianov, A., Gerasimov, E., Logacheva, M., and Penin, A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant J.* 88, 1058–1070. doi: 10.1111/tpj.13312

Lafos, M., Kroll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., and Schubert, D. (2011). Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. *PLoS Genet.* 7:e1002040. doi: 10.1371/journal.pgen.1002040

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

Larsson, J. (2020). *eulerr: Area-Proportional Euler and Venn Diagrams With Ellipses*. Available online at: https://cran.r-project.org/package=eulerr

Lee, Y., Rhee, J., Lee, S., Chung, G., Park, S., Segami, S., et al. (2018). Functionally redundant LNG3 and LNG4 genes regulate turgor-driven polar cell elongation through activation of XTH17 and XTH24. *Plant Mol. Biol.* 97, 23–36. doi: 10.1007/s11103-018-0722-0

Li, C., Gu, L., Gao, L., Chen, C., Wei, C., Qiu, Q., et al. (2016). Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in Arabidopsis. *Nat. Genet.* 48, 687–693. doi: 10.1038/ng.3555

Li, H., Ye, K., Shi, Y., Cheng, J., Zhang, X., and Yang, S. (2017). BZR1 positively regulates freezing tolerance via CBF-dependent and CBF-independent pathways in arabidopsis. *Mol. Plant.* 10, 545–559. doi: 10.1016/j.molp.2017.01.004

Li, R., Li, J., Li, S., Qin, G., Novák, O., Pěnčík, A., et al. (2014). ADP1 Affects plant architecture by regulating local auxin biosynthesis. *PLoS Genet.* 10:e1003954. doi: 10.1371/journal.pgen.1003954

Lu, F., Cui, X., Jenuwein, T., Cao, X., and Zhang, S. (2011). Arabidopsis REF6 is a histone H3 lysine 27 demethylase. *Nat. Genet.* 43, 715–719. doi: 10.1038/ng.854

Magwanga, R., Lu, P., Kirungu, J., Dong, Q., Hu, Y., Zhou, Z., et al. (2018). Cotton late embryogenesis abundant (LEA2) genes promote root growth and confer drought stress tolerance in transgenic Arabidopsis thaliana. *Genes Genomes Genetics* 8, 2781–2803. doi: 10.1534/g3.118.200423

Mara, C., and Irish, V. (2008). Two GATA transcription factors are downstream effectors of floral homeotic gene action in Arabidopsis. *Plant Physiol.* 147, 707–718. doi: 10.1104/pp.107.115634

Marowa, P., Ding, A., and Kong, Y. (2016). Expansins: roles in plant growth and potential applications in crop improvement. *Plant Cell Rep.* 35, 949–965. doi: 10.1007/s00299-016-1948-4

McClung, C., and Gutiérrez, R. (2010). Network news: prime time for systems biology of the plant circadian clock. *Curr. Opin. Genet. Dev.* 20, 588–598. doi: 10.1016/j.gde.2010.08.010

Meister, R., Oldenhof, H., Bowman, J., and Gasser, C. (2005). Multiple protein regions contribute to differential activities of YABBY proteins in reproductive development. *Plant Physiol.* 137, 651–662. doi: 10.1104/pp.104.055368

Nietzsche, M., Schießl, I., and Börnke, F. (2014). The complex becomes more complex: protein-protein interactions of SnRK1 with DUF581 family proteins provide a framework for cell- and stimulus type-specific SnRK1 signaling in plants. *Front. Plant Sci.* 5:54. doi: 10.3389/fpls.2014.00054

Pfluger, J., and Wagner, D. (2007). Histone modifications and dynamic regulation of genome accessibility in plants. *Curr. Opin. Plant Biol.* 10, 645–652. doi: 10.1016/j.pbi.2007.07.013

Plackett, A., Powers, S., Phillips, A., Wilson, Z., Hedden, P., and Thomas, S. (2017). The early inflorescence of Arabidopsis thaliana demonstrates positional effects in floral organ growth and meristem patterning. *Plant Reprod.* 31, 171–191. doi: 10.1007/s00497-017-0320-3

Qi, T., Huang, H., Song, S., and Xie, D. (2015). Regulation of jasmonate-mediated stamen development and seed production by a bHLH-MYB complex in Arabidopsis. *Plant Cell* 27, 1620–1633. doi: 10.1105/tpc.15.00116

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna.

Reeves, P., Ellis, C., Ploense, S., Wu, M., yadav, v., tholl, d., et al. (2012). a regulatory Network for Coordinated Flower Maturation. *PLoS Genet.* 8:e1002506. doi: 10.1371/journal.pgen.1002506

Ripoll, J., Zhu, M., Brocke, S., Hon, C., Yanofsky, M., Boudaoud, A., et al. (2019). Growth dynamics of the Arabidopsis fruit is mediated by cell expansion. *Proc. Natl. Acad. Sci. U.S.A.* 116, 25333–25342. doi: 10.1073/pnas.1914096116

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., et al. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinform.* 18:529. doi: 10.1186/s12859-017-1934-z

Seifert, G., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161. doi: 10.1146/annurev.arplant.58.032806.103801

Smyth, D., Bowman, J., and Meyerowitz, E. (1990). Early flower development in Arabidopsis. *Plant Cell* 2:755. doi: 10.2307/3869174

Song, S., Qi, T., Huang, H., Ren, Q., Wu, D., Chang, C., et al. (2011). The jasmonate-ZIM domain proteins interact with the R2R3-MYB transcription factors MYB21 and MYB24 to affect jasmonate-regulated stamen development in arabidopsis. *Plant Cell* 23, 1000–1013. doi: 10.1105/tpc.111.083089

Spartz, A., Ren, H., Park, M., Grandt, K., Lee, S., Murphy, A., et al. (2014). SAUR inhibition of PP2C-D phosphatases activates plasma membrane H+-ATPases to promote cell expansion in Arabidopsis. *Plant Cell* 26, 2129–2142. doi: 10.1105/tpc.114.126037

Stebbins, G. L. (1974). *Flowering Plants: Evolution Above the Species Level*. Cambridge, MA: Harvard University Press.

Tarazona, S., Garcia-Alcalde, F., Dopazo, J., Ferrer, A., and Conesa, A. (2011). Differential expression in RNA-seq: a matter of depth. *Genome Res.* 21:4436. doi: 10.1101/gr.124321.111

Van Leene, J., Hollunder, J., Eeckhout, D., Persiau, G., Van De Slijke, E., Stals, H., et al. (2010). Targeted interactomics reveals a complex core cell cycle machinery in Arabidopsis thaliana. *Mol. Syst. Biol.* 6:397. doi: 10.1038/msb.2010.53

Wager, A., and Browse, J. (2012). Social network: JAZ protein interactions expand our knowledge of jasmonate signaling. *Front. Plant Sci.* 3:41. doi: 10.3389/fpls.2012.00041

Wang, H., Liu, C., Cheng, J., Liu, J., Zhang, L., He, C., et al. (2016). Arabidopsis flower and embryo developmental genes are repressed in seedlings by different combinations of polycomb group proteins in association with distinct sets of Cis-regulatory elements. *PLoS Genet.* 12:e1005771. doi: 10.1371/journal.pgen.1005771

Wang, R., Liu, X., Liang, S., Ge, Q., Li, Y., Shao, J., et al. (2015). A subgroup of MATE transporter genes regulates hypocotyl cell elongation in Arabidopsis. *J. Exp. Bot.* 66, 6327–6343. doi: 10.1093/jxb/erv344

Wellmer, F., Bowman, J., Davies, B., Ferrándiz, C., Fletcher, J., Franks, R., et al. (2013). Flower development: open questions and future directions. *Methods Mol. Biol.* 1110, 103–124. doi: 10.1007/978-1-4614-9408-9_5

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Softw.* 4:1686. doi: 10.21105/joss.01686

Wright, S., Kalisz, S., and Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Proc. R. Soc. B: Biol. Sci.* 280:20130133. doi: 10.1098/rspb.2013.0133

Yan, W., Chen, D., Smaczniak, C., Engelhorn, J., Liu, H., Yang, W., et al. (2018). Dynamic and spatial restriction of Polycomb activity by plant histone demethylases. *Nat. Plants.* 4, 681–689. doi: 10.1038/s41477-018-0219-5

Yang, L., Wang, Z., and Hua, J. (2019). "Measuring cell ploidy level in arabidopsis thaliana by flow cytometry," in: *Plant Innate Immunity. Methods in Molecular Biology*. ed W. Gassmann W (New York, NY: Humana). doi: 10.1007/978-1-4939-9458-8_11

Yu, X., Li, L., Li, L., Guo, M., Chory, J., and Yin, Y. (2008). Modulation of brassinosteroid-regulated gene expression by jumonji domain-containing proteins ELF6 and REF6 in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7618–7623. doi: 10.1073/pnas.0802254105

Yu, X., Li, L., Zola, J., Aluru, M., Ye, H., Foudree, A., et al. (2011). A brassinosteroid transcriptional network revealed by genome-wide identification of BESI target genes in Arabidopsis thaliana. *Plant J.* 65, 634–646. doi: 10.1111/j.1365-313X.2010.04449.x

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., et al. (2007). Whole genome analysis of histone H3 Lysine 27 trimethylation in Arabidopsis. *PLoS Biol.* 5:e129. doi: 10.1371/journal.pbio.0050129

Zheng, S., Hu, H., Ren, H., Yang, Z., Qiu, Q., Qi, W., et al. (2019). The Arabidopsis H3K27me3 demethylase JUMONJI

13 is a temperature and photoperiod dependent flowering repressor. *Nat. Commun.* 10, 1303–1311. doi: 10.1038/s41467-019-09310-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 in Plant Science

REVIEW
published: 25 February 2021
doi: 10.3389/fpls.2021.637244

The Epigenetic Faces of ULTRAPETALA1

*Diego Ornelas-Ayala[1], Adriana Garay-Arroyo[1,2], Berenice García-Ponce[1], Elena R. Álvarez-Buylla[1,2] and María de la Paz Sanchez[1]**

[1] Laboratorio de Genética Molecular, Epigenética, Desarrollo y Evolución de Plantas, Instituto de Ecología, Universidad Nacional Autónoma de México, 3er Circuito Ext. Junto a J. Botánico, Ciudad Universitaria, UNAM, Mexico City, Mexico, [2] Centro de Ciencias de la Complejidad (C3), Universidad Nacional Autónoma de México, Mexico City, Mexico

ULTRAPETALA1 (ULT1) is a versatile plant-exclusive protein, initially described as a trithorax group (TrxG) factor that regulates transcriptional activation and counteracts polycomb group (PcG) repressor function. As part of TrxG, ULT1 interacts with ARABIDOPSIS TRITHORAX1 (ATX1) to regulate H3K4me3 activation mark deposition. However, our recent studies indicate that ULT1 can also act independently of ATX1. Moreover, the ULT1 ability to interact with transcription factors (TFs) and PcG proteins indicates that it is a versatile protein with other roles. Therefore, in this work we revised recent information about the function of Arabidopsis ULT1 to understand the roles of ULT1 in plant development. Furthermore, we discuss the molecular mechanisms of ULT1, highlighting its epigenetic role, in which ULT1 seems to have characteristics of an epigenetic molecular switch that regulates repression and activation processes via TrxG and PcG complexes.

Keywords: ULTRAPETALA1, TrxG, PcG, ATX1, Molecular epigenetic switch, Arabidopsis

OPEN ACCESS

Edited by:
Gang Wu,
Zhejiang Agriculture and Forestry University, China

Reviewed by:
Chris Helliwell,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia
Li Pu,
Chinese Academy of Agricultural Sciences, China

****Correspondence:***
María de la Paz Sanchez
mpsanchez@iecologia.unam.mx

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

Received: *03 December 2020*
Accepted: *21 January 2021*
Published: *25 February 2021*

Citation:
Ornelas-Ayala D, Garay-Arroyo A, García-Ponce B, R Álvarez-Buylla E and Sanchez MP (2021) The Epigenetic Faces of ULTRAPETALA1. Front. Plant Sci. 12:637244. doi: 10.3389/fpls.2021.637244

INTRODUCTION

In multicellular organisms, epigenetic regulation plays crucial roles for the correct deployment of developmental programs and for the establishment of cell fates. Epigenetic mechanisms include post-translational histone modifications (PHM) that modulate chromatin structure to regulate gene expression. The trithorax group (TrxG) is an epigenetic protein complex able to regulate transcriptional activation through trimethylation of lysine 4 and 36 of histone H3 (H3K4me3 and H3K36me3) as well as other associated PHMs (Schuettengruber et al., 2011). TrxG proteins are those that belong to complexes counteracting of polycomb group (PcG) repressive activity at the same set of target genes (Grimaud et al., 2006); however, other proteins that act together with TrxG on PcG or non-PcG target genes are also considered TrxG (Schuettengruber et al., 2007).

In plants, TrxG participates in different developmental processes from embryogenesis to floral development, regulating gene expression of several transcription factors (TFs) involved in stem cell maintenance, cell fate identity, and cell proliferation and differentiation (Sanchez et al., 2015; Fletcher, 2017). The plant TrxG complex has been identified by homology to known TrxG proteins in animals or by genetic characterization based on their ability to counteract PcG mutant phenotypes (Fletcher, 2017). In this regard, SET histone methyltransferases (HMTs) of MLL and SET families, COMPASS-like proteins such as WDR5, ASH2L and RBBP5, and ATP-dependent chromatin-remodeling factors such as BRM, CHD and BPTF, have been described in plants

(Avramova, 2009; Schuettengruber et al., 2011; Sanchez et al., 2015) (**Figure 1**). In *Arabidopsis thaliana* (hereafter Arabidopsis), the main HMTs of TrxG that catalyze the H3K4me3 mark are the ARABIDOPSIS TRITHORAX1 (ATX1) and the ARABIDOPSIS TRITHORAX-RELATED 3/SETDOMAIN GROUP 2 (ATXR3/SDG2) (Alvarez-Venegas et al., 2003; Berr et al., 2010; Guo et al., 2010; Chen et al., 2017), although until now, only ATX1 has been found to form a complex within the core of Arabidopsis COMPASS-like complex described (Jiang et al., 2011). Interestingly, it has been reported that the plant TrxG group includes a unique protein named ULTRAPETALA1 (ULT1) (**Figure 1**), whose structure differs from all TrxG components reported in animals and yeast. ULT1 has been defined as a TrxG factor by counteract PcG silencing and by its physical interaction with ATX1 (Carles and Fletcher, 2009; Pu et al., 2013). However, our recent study indicates that ULT1 can act independently of ATX1, in a tissue-specific fashion (Ornelas-Ayala et al., 2020). Moreover, the interactions of ULT1 with PcG proteins (Xu et al., 2018) suggest other roles of ULT1 as well. Therefore, here we review recent information on the structure of the ULT1 protein, its interactions with other proteins, and its gene targets, as well as the phenotypic analysis of loss-of-function mutants to understand the roles of ULT1 in plant development. Furthermore, we discuss the molecular mechanisms in which ULT1 is involved, as well as its possible function as an epigenetic molecular switch that regulates repression and activation processes via TrxG and PcG complexes.

WHAT THE ULT1 STRUCTURE REVEALS ABOUT ITS FUNCTION

In Arabidopsis, ULT1 has been described as a SAND (named after Sp100, AIRE, NucP41/75, DEAF-1) domain protein that also contains a B-box motif (**Figure 2A**), a motif that seems to be important for protein-protein interaction (Torok and Etkin, 2001; Carles et al., 2005; Khanna et al., 2009). In the case of OsULT1 from *Oryza sativa*, it has been shown that is important for its multimerization (Roy et al., 2019). Meanwhile, the SAND domain has a DNA-binding function (Bottomley et al., 2001), and it is conserved in plants and animals in vast combinations with other protein domains on the Viridiplantae and metazoan lineages. The Clorophyta lineage contains a single-SAND domain protein RegA, whereas in the Embryophyte lineage only ULT and ATX3 (ARABIDOPSIS THRITHORAX3) proteins and its paralogs contain a SAND domain (Kirk et al., 1999; Nedelcu, 2019). In ULT proteins, the SAND domain is unique, whereas in ATX3, it appears in combination with the SET-like and PHD domains (Nedelcu, 2019). The SAND domain in combination with other protein domains has also been related to chromatin interactions and transcriptional regulation. For instance, AIRE (Autoimmune Regulator) is capable of interacting with chromatin through its PHD domain. AIRE binds specifically unmethylated H3K4 residues and it is proposed that this binding is important for its function as a transcriptional activator (Org et al., 2008). Moreover, the AIRE protein can associate with DNA transcriptional control elements and factors involved in pre-mRNA processing (Abramson et al., 2010) and also can be acetylated by the CBP (CREB Binding Protein) and the p300 histone acetyltransferases to enhance its transactivation activity (Saare et al., 2012). Therefore, the SAND domain is a DNA-binding module characteristic of chromatin-dependent transcriptional regulation. In fact, by *in vitro* assays, it has been shown that the SAND domain of human DEAF-1 (Deformed Epidermal Autoregulatory Factor-1) homolog recognizes the 5′-TTCG-3′ sequence (Bottomley et al., 2001). This sequence differs from what has been reported in plants, where the SAND domain of recombinant OsULT1, has affinity for the 5′-GAGAG-3′ sequence (Roy et al., 2019).

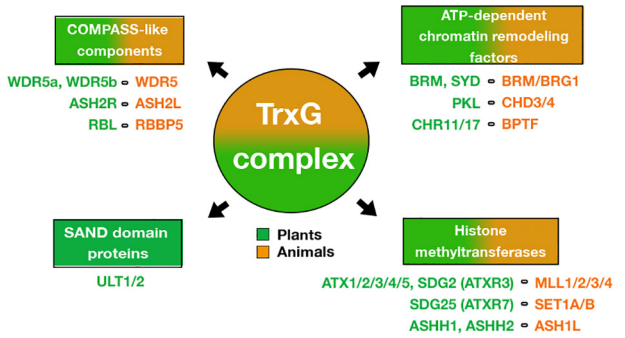

FIGURE 1 | Factors of the TrxG complex. Different proteins of the TrxG animal complex that are conserved in plants such as (i) ATP-dependent chromatin-remodeling factors, (ii) SET domain-containing proteins that catalyze histone methylation, and (iii) COMPASS-like proteins. It is noteworthy that plant TrxG includes SAND-domain proteins that are not conserved in animals. The list of names below of each TrxG component shows the proteins described in Arabidopsis (green color) and their mammalian counterparts (orange color).

Most of the SAND domain proteins of the different lineages are involved in developmental processes such as cell proliferation, cell differentiation, tissue homeostasis and organ formation (Nedelcu, 2019). For instance, in the multicellular green alga *Volvox carteri*, RegA is involved in somatic cell differentiation (Kirk et al., 1999), while the DEAF-1 protein is necessary for embryonic development in *Drosophila melanogaster* (Veraksa et al., 2002), and its ortholog in mammals is involved in breast epithelial cell differentiation (Barker et al., 2008). In addition, AIRE is an important transcriptional activator to regulate autoimmune processes in the thymus (Abramson et al., 2010).

In plants, ULT1 functions have been described only for Arabidopsis and rice (see below); however, several ULT1 sequences have been reported in other species. In this kingdom, ULT1 seems to be a protein exclusive to Angiosperms, since Gymnosperm, Lycophytes or Mosses lack sequences homologous to ULT1. In angiosperms ULT1 is highly conserved in different species of Eudicots, Monocotyledons, and even in Amborellales, considered one of the most basal angiosperms (Chase et al., 2016), the latter being closer to Eudicots than

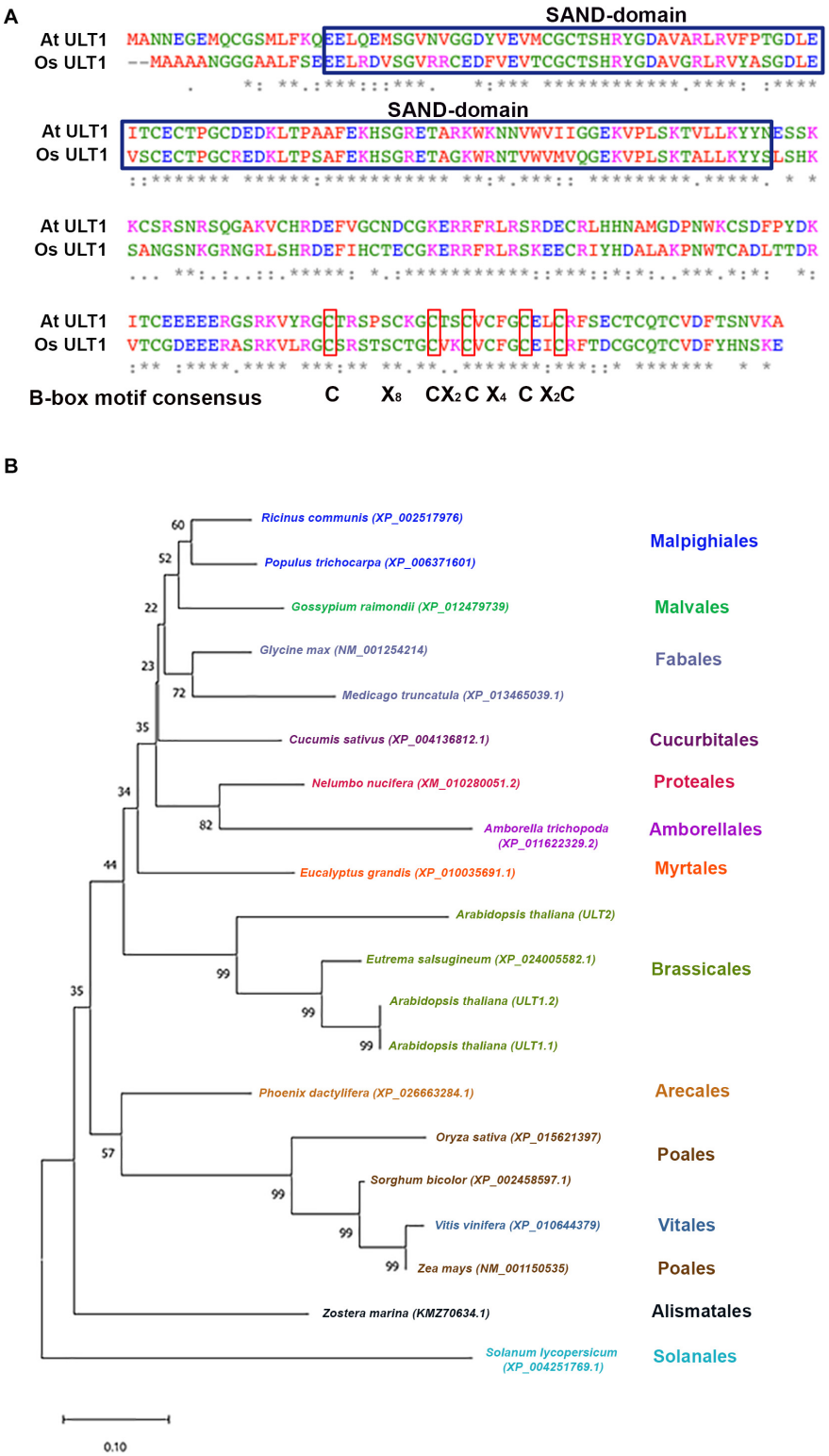

FIGURE 2 | ULTRAPETALA1 is conserved in different angiosperm species. **(A)** The SAND-domain and B-box motif of ULT1 from *Arabidopsis thaliana* and its alignment with ULT1 from *Oryza sativa*. The asterisks show identical residues; colons (:) and periods (.) show residues with strongly and weakly similar properties, respectively. The blue boxes show the SAND-domain and the red boxes represent the B-box consensus motif. **(B)** Phylogenetic analysis generated using the neighbor-joining method based on the ULT1 protein sequence of selected plant species. Numbers at nodes represent bootstrap percentages based on 10,000 samplings. The scale bars represent 0.1 substitutions per site.

to Monocotyledons (**Figure 2B**). The topology of neighbor-joining phylogenetic analysis shows a clear clade distribution according to plant orders, with the exception of *Vitis vinifera* that is closer to Poales (**Figure 2B**). Evolutionary conservation is also observed for Arabidopsis ULT2, a paralog of ULT1, which conserved a similar protein structure that includes the SAND domain (Carles et al., 2005). The high identity of ULT1 proteins in these species predicts similar functions among them.

THE ROLE OF ULT1 AS PART OF TrxG EPIGENETIC COMPLEX

The first reports on ULT1 function were made by analyzing the *ULT1* loss and gain-of-function mutant plants (Fletcher, 2001; Carles et al., 2004, 2005; Carles and Fletcher, 2009). Indeed, loss of function of ULT1 delays differentiation and increases shoot and floral meristem size, producing extra-floral organs such as sepals and petals, hence the name ULTRAPETALA (Fletcher, 2001; Carles et al., 2004). In the shoot apical meristem (SAM), ULT1 positively regulates the expression of *APETALA3 (AP3)* and *AGAMOUS (AG)* (**Figure 3A**), two genes of the ABC flower organ identity model (Carles and Fletcher, 2009). However, ULT1 was also described as a negative regulator of *WUSCHEL (WUS)* expression (**Figure 3B**), a TF that maintains stem cells in the meristems and must be repressed in order to establish floral determinacy (Carles et al., 2004). Therefore, these reports describe ULT1 as a putative transcriptional regulator, involved in shoot meristem maintenance and floral meristem differentiation and determinacy. Nevertheless, the opposite regulation between ULT1 and CURLY LEAF (CLF), an HMT of the Arabidopsis PcG repressive complex, observed in some vegetative and reproductive organs (Carles and Fletcher, 2009), as well as the antagonistic function of ULT1 with EMBRYONIC FLOWER1 (EMF1), another PcG component (Pu et al., 2013), together with the ability of ULT1 to physically interact with the ATX1, have led to propose ULT1 as a TrxG factor with coactivator

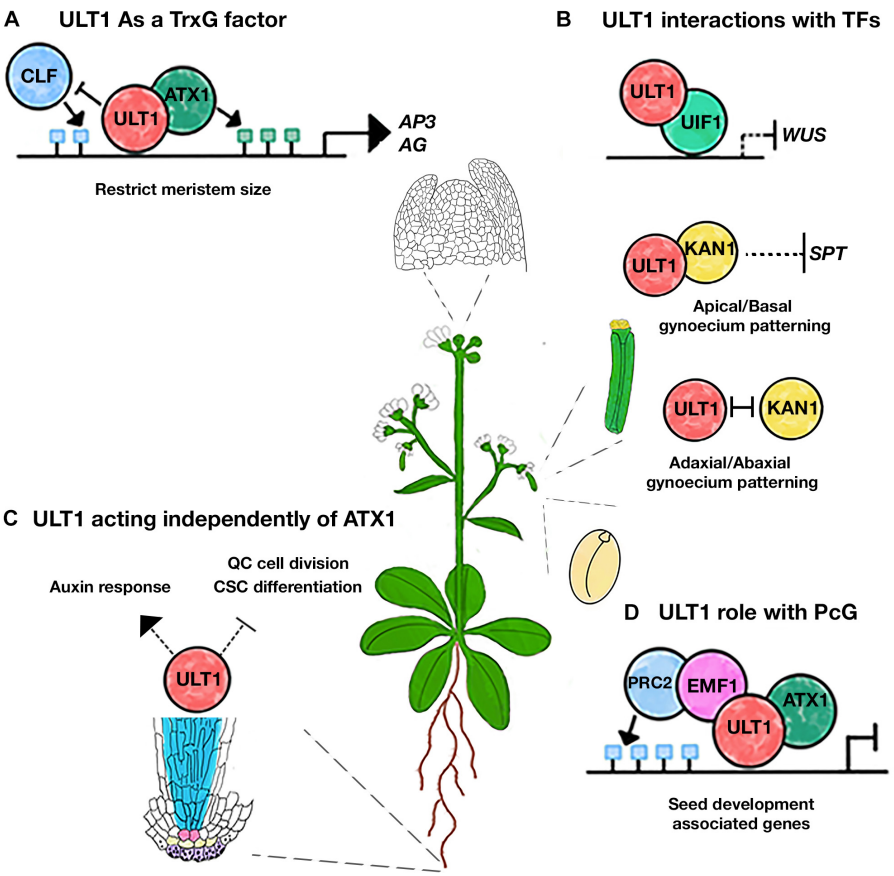

FIGURE 3 | Different roles of ULT1 during Arabidopsis development. **(A)** ULT1 interacts with ATX1 to counteract PcG functions in the SAM. **(B)** ULT1 interacts with UIF1 to repress the *WUS* expression in the SAM. ULT1 may also act with KAN1 to regulate the apical/basal patterning of the gynoecium and it can function antagonistically with KAN1 to regulate the adaxial/abaxial patterning of the gynoecium. **(C)** ULT1 regulates auxin response, the QC cell division rate, and the columella stem cells (CSC) differentiation to maintain root SCN, independently of ATX1. **(D)** ULT1 interacts with EMF1 to keep the repression of seed development genes by maintaining the H3K27me3 marks. The green and blue boxes represent H3K4me3 and H3K27me3 marks, respectively.

properties of some genes related to the SAM development (Carles and Fletcher, 2009).

Furthermore, despite the lack of HMT activity of ULT1, it has been suggested that *ult1* mutant plants have lower levels of H3K4me3 marks on *AG* and *AP3* genes, which are associated with an increase of H3K27me3 PcG mark on these ULT1 targets (Carles and Fletcher, 2009; Pu et al., 2013), evidencing the ability of ULT1 to regulate these epigenetic marks. Interestingly, the 5′-GAGAG-3′ Arabidopsis PRE motifs recognized by CLF and its functional homolog SWINGER (SWN), as well as by other core components of PcG (Deng et al., 2013; Xiao et al., 2017; Shu et al., 2019), can also be recognized by the OsULT1 SAND-domain (Roy et al., 2019). Given that the ULT1 SAND-domains from rice and Arabidopsis share 90.91% of similarity (**Figure 2A**), it could be predicted that Arabidopsis ULT1 can bind through its SAND domain to the same sites as PcG proteins and thereby interfere with H3K27me3 marks.

All of these reports indicate that ULT1 is a unique SAND-domain protein that is part of a TrxG complex; neither in animals nor in yeast is there evidence of SAND-domain proteins in the TrxG complexes described so far.

DIFFERENT TISSUES, DIFFERENT ULT1 MECHANISMS

Although ULT1 is able to bind to ATX1, its interactions with other TrxG components are unclear. Unlike the other members of the TrxG, ULT1 has a very discrete expression pattern, being mainly expressed in young organ primordia and shoot and root meristems (Carles et al., 2005; Ornelas-Ayala et al., 2020). This suggests that ULT1 has a tissue-specific regulation rather than a general expression pattern as do the other TrxG members.

Genome-wide analyses have revealed that ULT1-regulated genes are involved in different developmental processes (Tyler et al., 2019). Besides its function in SAM development (Fletcher, 2001), ULT1 participates in different stress processes (Pu et al., 2013; Tyler et al., 2019). In addition, we recently found that ULT1 is necessary for root stem cell niche (SCN) maintenance (**Figure 3C**), including the cell division rate of the Quiescent Center (QC) and the undifferentiated state of the columella stem cells (Ornelas-Ayala et al., 2020). Interestingly and in contrast to its role in the SAM, our genetic analyses of *atx1* and *ult1* single and double mutants revealed that in the root apical meristem (RAM) ULT1 acts independently of ATX1 (Ornelas-Ayala et al., 2020). The *ult1* mutants showed a diminished response to auxins, demonstrated by a down regulation of some efflux *PIN* transporter genes and the *DR5-GUS* reporter, as well as a premature columella stem cell differentiation (Ornelas-Ayala et al., 2020). Contrary to this, *atx1* mutants do not seem to have defects in auxin response, whereas the columella stem cell differentiation seems to be delayed; besides, in contrast to *atx1* mutants, *ult1* plants did not show any changes in the root and RAM length (Napsucialy-Mendivil et al., 2014; Ornelas-Ayala et al., 2020).

Although the studies of the relationship between ULT1 and ATX1 in the SAM were carried out by single mutant analysis and biochemical methods, and in the RAM were carried out by genetic analysis of double mutants, with these studies, it is possible to establish that ULT1 can act by different mechanisms in the SAM and in the RAM, one of which requires ATX1 to regulate some aspects of floral development while in the other, ULT1 maintains SCN homeostasis in ATX1-independent manner.

In this regard, 18.7% (2859) of Arabidopsis genes are deregulated in *atx1* loss-of-function mutants, whereas 5.6% (856) are deregulated in *ult1* mutants, and among them only a little subset is shared (1.1%; 170 genes) in both *atx1* and *ult1* mutants (Xu et al., 2018); although this does not mean that it is a direct regulation by ATX1 or ULT1, it reflects the behavior of genes that do not always act together. In fact, by ChIP-seq analysis, it has also been determined that out of the 2,276 Arabidopsis TFs annotated (Perez-Rodriguez et al., 2010; Jin et al., 2017), ATX1 is bound to 43 (1.88%) of these, whereas ULT1 to 67 (2.9%) and only in 18 (0.8%) of these are bound both ATX1 and ULT1 (Xu et al., 2018), evidencing that ATX1 and ULT1 have independent targets.

The ATX1-independent function of ULT1 raises the question whether ULT1 acts together with other HMTs of the TrxG complex or by a TrxG-independent mechanism or both in different developmental processes. The analysis of ULT1 protein interactions in different developmental contexts could provide evidence compatible with both mechanisms as shown below.

ULT1 ACTS TOGETHER WITH SOME TRANSCRIPTION FACTORS

The presence of the B-box motif in ULT1 suggests multiple interactions with other proteins. Indeed, ULT1 interacts with some TFs (**Figure 3B**). One of these is the GARP family transcription factor KANADI1 (KAN1), described as a transcriptional repressor, involved in the patterning of the abaxial polarity of leaves and the gynoecium (Eshed et al., 2001; Pires et al., 2014; Xie et al., 2015). ULT1 interacts physically with KAN1 and genetic analysis indicates that they participate together in the apical-basal polarity of the gynoecium, restricting the *SPATULA (SPA)* expression, which promotes carpel marginal tissue apical style and stigma tissue formation (**Figure 3B**). But also, ULT1 and KAN1 may act antagonistically to regulate the adaxial-abaxial axis of the gynoecium (Pires et al., 2014; **Figure 3B**). ULT2 also physically interacts with KAN1, performing redundant roles on the apical-basal gynoecium patterning (Monfared et al., 2013; Pires et al., 2014).

Furthermore, the physical interaction of ULT1 with the MYB domain-containing TF ULTRAPETALA INTERACTING FACTOR 1 (UIF1) has been reported. UIF1 binds to *WUS* and *AG* regulatory sequences in the floral meristem (Moreau et al., 2016). Given that UIF1 acts as a transcriptional repressor, it has been suggested that it represses *WUS* expression when interacting with ULT1, to establish floral meristem determinacy (Moreau et al., 2016; **Figure 3B**).

These reports have led to suggestions that ULT1 can act as a link between chromatin-remodeling factors and some TFs (Pires et al., 2014). However, other evidence will be needed to indicate whether the combined function of ULT1 with

these TFs depends on the other components of TrxG or is TrxG-independent.

CAN ULT1 ACT IN DIFFERENT TrxG COMPLEXES?

The lower levels of H3K4me3 marks detected in some genes in the *ult1* mutants compared with those observed in *atx1* mutants (Xu et al., 2018) support the idea that ULT1 can act together with TrxG complex but independently of ATX1, suggesting the existence of different TrxG complexes, through which ULT1 can perform its function. In this regard, multiple SET or MLL HMT homologues from yeasts and animals that can form different COMPASS-like complexes and predict the existence of different TrxG complexes in plants (Schuettengruber et al., 2011). The Arabidopsis compass-like complex reported so far contains ATX1 as the H3K4me3 HMT (Jiang et al., 2009, 2011); however, there are other HMTs of H3K4, such as ATX1/SDG27, ATX2/SDG30, ATXR3/SDG2 and ATXR7/SDG25, that could form different COMPASS complexes (Sanchez et al., 2015). Indeed, it has been demonstrated that the SAND domain of OsULT1 is responsible for interacting with the SET-domain of OsTRX1, an ATX1 ortholog (Roy et al., 2019). The high similarity of Arabidopsis and rice SAND-domains of ULT1 (**Figure 2A**) suggests that ULT1 can also interact with different proteins with a SET-domain.

Of particular interest is ATXR3/SDG2, reported as the main HMT of the Arabidopsis (Guo et al., 2010). ATXR3/SDG2 does not have a significant sequence homology with other SDGs outside of the SET domain. However, the gene encoding this protein is broadly expressed and is crucial for multiple Arabidopsis developmental processes, regulating 46.4% of all H3K4me3 sites in the Arabidopsis genome (Berr et al., 2010; Guo et al., 2010; Chen et al., 2017). In root tissues, the *sdg2* loss-of-function mutant shares some phenotypes with *ult1* mutants, such as disorganization of the SCN, early differentiation of the columella stem cells, and diminished auxin response (Yao et al., 2013; Ornelas-Ayala et al., 2020). Although it is still unknown whether ULT1 interacts with SDG2, the similarities in their phenotypes raises the possibility that ULT1 could act with SDG2 in some developmental contexts.

DOES ULT1 FUNCTION AS A MOLECULAR EPIGENETIC SWITCH?

Besides the interactions with TFs and TrxG factors, ULT1 also interacts with EMF1 (Xu et al., 2018). EMF1 is the plant-specific protein proposed as a component of Polycomb repressive complex 1 (PRC1), acting as a bridge to the Polycomb repressive complex 2 (PRC2) (Calonje et al., 2008; Wang et al., 2014). Although the relevance of such interaction is unknown, the H3K27me3 abundancy on some EMF1-target genes associated with seed development decreases more in the *emf1/ult1/atx1* triple mutant than in *emf1*, *atx1,* or *ult1* single mutant (Xu et al., 2018). In this framework, it has been proposed that ULT1 interacts with ATX1 to form a complex with PRC2 through EMF1 to maintain the H3K27me3 marks and a chromatin repressive state (Xu et al., 2018). This model suggests that ULT1 not only acts to antagonize the PcG activity; instead, it could act together with PRC2, maintaining the repression states of some targets, through the maintenance of the H3K27me3 mark (**Figure 3D**). For instance, it has been seen that the *ult1* mutants have more upregulated genes than down-regulated genes (Xu et al., 2018; Tyler et al., 2019). Interestingly, the MADS-box *FLOWERING LOCUS C (FLC)* gene, which is activated by TrxG and repressed by PcG (Whittaker and Dean, 2017), is upregulated (∼4.35 fold) in *ult1* mutant plants (Pu et al., 2013; Xu et al., 2018; Tyler et al., 2019), contrary to what is expected for TrxG mutants. Besides, ULT1 binding to the *FLC* locus supports a direct regulation (Xu et al., 2018). Moreover, the *FLC* upregulation is higher in *ult1/emf1* double mutants than in the *emf1* single mutant (Pu et al., 2013). Hence, loss of *ULT1* function enhances *emf1* upregulation on *FLC*. In contrast, a different behavior was observed on genes that are positively regulated by ULT1, e.g., *AG*, whose upregulation in *emf1* loss-of-function mutants is abated in the double mutant *ult1/emf1* plants (Pu et al., 2013). Although additional experiments are needed, these observations support the involvement of ULT1 in transcriptional repression. Moreover, the repressive function of ULT1 could be compatible with *WUS* repression via UIF1 (Moreau et al., 2016), where PcG could also be participating, as it has been reported (Xu and Shen, 2008).

Given these observations, we suggest two modes of ULT1 action: one through TrxG to regulate transcriptional activation via H3K4me3 deposition, which can be ATX1 dependent or independent, and another, through PcG via EMF1 to repress transcription.

The apparent dual function of ULT1 has led us to wonder whether ULT1 can act as a molecular epigenetic switch, regulating transcriptional repression and activation via PcG and TrxG, respectively. The presence of molecular epigenetic switches allows a dynamic regulation, capable of changing gene expression quickly and efficiently to face different environmental and developmental states. The existence of bivalent chromatin domains provides persuasive evidence of molecular epigenetic switches that regulate gene expression (Hoffmann et al., 2015). The bivalent domains produced by TrxG and PcG serve to keep developmental genes on standby, primed for subsequent expression and to protect against unscheduled expression, reducing transcriptional noise in favor of robust developmental decisions (Hoffmann et al., 2015). Although in plant biological studies, bivalent marks in the same locus have been little addressed and still remain elusive, finding proteins involved in both activation and repression processes shows the relevance of bivalent marks to regulating gene expression quickly and efficiently. In this regard, ULT1 fulfills the main features to act as a molecular epigenetic switch: (i) interaction with both TrxG and PcG proteins, (ii) the ability to increase or decrease gene expression, and (iii) the ability to regulate the deposition of H3K4me3 and H3K27me3 marks. However, establishing whether these characteristics converge into specific genes in time and/or space is still necessary, in such a way that ULT1 can be a link to

load the TrxG or PcG complexes and consequently regulate gene expression accordingly.

CONCLUSION AND PERSPECTIVES

Current knowledge reveals ULT1 to be a versatile protein able to interact with TFs, TrxG, and PcG proteins to regulate gene expression of several developmental processes: (1) ULT1 activates genes related to floral development through its interaction with ATX1, (2) in association with UIF1, ULT1 represses *WUS* expression to regulate shoot and floral meristem homeostasis, (3) ULT1 is also involved in the regulation of gynoecium patterning, in which it interacts with KAN1 to repress *SPT*, (4) ULT1 together with EMF1 maintains repressive marks of some genes related to seed development, and (5) ULT1, independently of ATX1, is involved in the root SCN maintenance (**Figure 3**). The ability of ULT1 to regulate both gene expression and repression by modulation of H3K4me3 and H3K27me3 bivalent marks makes this protein a suitable candidate to regulate bivalent genes that can be in a poised state, waiting for future instructions from the cell. The role of ULT1, independent of ATX1 in roots tissues, suggests a function with other TrxG factors, evidencing the possible existence of different TrxG complexes that could be formed in a tissue-specific fashion in which ULT1 could be involved.

The complexity of ULT1 interactions, the phenotypes reported for *ult1* mutants, and their genome-wide effects make it difficult to define modes of action of ULT1. However, these reports illustrate four possible ways of action for ULT1: (i) together with TrxG factors, (ii) with PcG factors, (iii) outside of both TrxG/PcG complex, and (iv) in association with TFs. Furthermore, a possible mechanism cannot be ruled out through which ULT1 and TrxG or PcG converge in association with TFs.

It would be important to study specific ULT1 targets in different developmental and/or tissue-specific stages to analyze the ULT1 involvement on its activation or repression, which could shed light on the role of ULT1 in association with TrxG and PcG complexes, as a molecular epigenetic switch. Therefore, the future challenge is to define whether ULT1 acts by different mechanisms or in a single mechanism that involves all reported interactions. In this regard, additional research is needed to define whether these mechanisms can coexist or are tissue-, cell type-, or loci-specific.

AUTHOR CONTRIBUTIONS

DO-A and MPS conceived and wrote the review. AG-A, ERA-B, and BG-P wrote the review. All authors have read and approved this version of the manuscript.

FUNDING

This work was supported by UNAM-DGAPA-PAPIIT IN203220, IN206220, IN200920, and IN211721. CONACyT 102987 and 102959.

ACKNOWLEDGMENTS

We thank Diana B. Sánchez Rodríguez for her logistical support. Diego Ornelas-Ayala is a Ph.D. student from the Posgrado en Ciencias Biomédicas, Universidad Nacional Autónoma de México, Mexico, and recipient of a fellowship from CONACyT (588728), Mexico.

REFERENCES

Abramson, J., Giraud, M., Benoist, C., and Mathis, D. (2010). Aire's partners in the molecular control of immunological tolerance. *Cell* 140, 123–135. doi: 10.1016/j.cell.2009.12.030

Alvarez-Venegas, R., Pien, S., Sadder, M., Witmer, X., Grossniklaus, U., and Avramova, Z. (2003). ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. *Curr. Biol.* 13, 627–637. doi: 10.1016/s0960-9822(03)00243-4

Avramova, Z. (2009). Evolution and pleiotropy of TRITHORAX function in *Arabidopsis*. *Int. J. Dev. Biol.* 53, 371–381. doi: 10.1387/ijdb.082664za

Barker, H. E., Smyth, G. K., Wettenhall, J., Ward, T. A., Bath, M. L., Lindeman, G. J., et al. (2008). Deaf-1 regulates epithelial cell proliferation and side-branching in the mammary gland. *BMC Dev. Biol.* 8:94. doi: 10.1186/1471-213X-8-94

Berr, A., McCallum, E. J., Menard, R., Meyer, D., Fuchs, J., Dong, A., et al. (2010). *Arabidopsis* SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell* 22, 3232–3248. doi: 10.1105/tpc.110.079962

Bottomley, M. J., Collard, M. W., Huggenvik, J. I., Liu, Z., Gibson, T. J., and Sattler, M. (2001). The SAND domain structure defines a novel DNA-binding fold in transcriptional regulation. *Nat. Struct. Biol.* 8, 626–633. doi: 10.1038/89675

Calonje, M., Sanchez, R., Chen, L., and Sung, Z. R. (2008). EMBRYONIC FLOWER1 participates in polycomb group-mediated AG gene silencing in *Arabidopsis*. *Plant Cell* 20, 277–291. doi: 10.1105/tpc.106.049957

Carles, C. C., Choffnes-Inada, D., Reville, K., Lertpiriyapong, K., and Fletcher, J. C. (2005). ULTRAPETALA1 encodes a SAND domain putative transcriptional regulator that controls shoot and floral meristem activity in *Arabidopsis*. *Development* 132, 897–911. doi: 10.1242/dev.01642

Carles, C. C., and Fletcher, J. C. (2009). The SAND domain protein ULTRAPETALA1 acts as a trithorax group factor to regulate cell fate in plants. *Genes Dev.* 23, 2723–2728. doi: 10.1101/gad.1812609

Carles, C. C., Lertpiriyapong, K., Reville, K., and Fletcher, J. C. (2004). The ULTRAPETALA1 gene functions early in Arabidopsis development to restrict shoot apical meristem activity and acts through WUSCHEL to regulate floral meristem determinacy. *Genetics* 167, 1893–1903. doi: 10.1534/genetics.104.028787

Chase, M., Christenhusz, M., Fay, M., Byng, J., Judd, W., Soltis, D., et al. (2016). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* 181:20.

Chen, L. Q., Luo, J. H., Cui, Z. H., Xue, M., Wang, L., Zhang, X. Y., et al. (2017). ATX3, ATX4, and ATX5 encode putative H3K4 methyltransferases and are critical for plant development. *Plant Physiol.* 174, 1795–1806. doi: 10.1104/pp.16.01944

Deng, W., Buzas, D. M., Ying, H., Robertson, M., Taylor, J., Peacock, W. J., et al. (2013). *Arabidopsis* polycomb repressive complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes. *BMC Genomics* 14:593. doi: 10.1186/1471-2164-14-593

Eshed, Y., Baum, S. F., Perea, J. V., and Bowman, J. L. (2001). Establishment of polarity in lateral organs of plants. *Curr. Biol.* 11, 1251–1260. doi: 10.1016/s0960-9822(01)00392-x

Fletcher, J. C. (2001). The ULTRAPETALA gene controls shoot and floral meristem size in *Arabidopsis*. *Development* 128, 1323–1333.

Fletcher, J. C. (2017). State of the Art: trxG factor regulation of post-embryonic plant development. *Front. Plant Sci.* 8:1925. doi: 10.3389/fpls.2017.01925

Grimaud, C., Negre, N., and Cavalli, G. (2006). From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. *Chromosome Res.* 14, 363–375. doi: 10.1007/s10577-006-1069-y

Guo, L., Yu, Y., Law, J. A., and Zhang, X. (2010). SET DOMAIN GROUP2 is the major histone H3 lysine [corrected] 4 trimethyltransferase in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18557–18562. doi: 10.1073/pnas.1010478107

Hoffmann, A., Zimmermann, C. A., and Spengler, D. (2015). Molecular epigenetic switches in neurodevelopment in health and disease. *Front. Behav. Neurosci.* 9:120. doi: 10.3389/fnbeh.2015.00120

Jiang, D., Gu, X., and He, Y. (2009). Establishment of the winter-annual growth habit via FRIGIDA-mediated histone methylation at FLOWERING LOCUS C in *Arabidopsis*. *Plant Cell* 21, 1733–1746. doi: 10.1105/tpc.109.067967

Jiang, D., Kong, N. C., Gu, X., Li, Z., and He, Y. (2011). Arabidopsis COMPASS-like complexes mediate histone H3 lysine-4 trimethylation to control floral transition and plant development. *PLoS Genet.* 7:e1001330. doi: 10.1371/journal.pgen.1001330

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Khanna, R., Kronmiller, B., Maszle, D. R., Coupland, G., Holm, M., Mizuno, T., et al. (2009). The *Arabidopsis* B-box zinc finger family. *Plant Cell* 21, 3416–3420. doi: 10.1105/tpc.109.069088

Kirk, M. M., Stark, K., Miller, S. M., Muller, W., Taillon, B. E., Gruber, H., et al. (1999). regA, a Volvox gene that plays a central role in germ-soma differentiation, encodes a novel regulatory protein. *Development* 126, 639–647.

Monfared, M. M., Carles, C. C., Rossignol, P., Pires, H. R., and Fletcher, J. C. (2013). The ULT1 and ULT2 trxG genes play overlapping roles in *Arabidopsis* development and gene regulation. *Mol. Plant* 6, 1564–1579. doi: 10.1093/mp/sst041

Moreau, F., Thevenon, E., Blanvillain, R., Lopez-Vidriero, I., Franco-Zorrilla, J. M., Dumas, R., et al. (2016). The Myb-domain protein ULTRAPETALA1 INTERACTING FACTOR 1 controls floral meristem activities in *Arabidopsis*. *Development* 143, 1108–1119. doi: 10.1242/dev.127365

Napsucialy-Mendivil, S., Alvarez-Venegas, R., Shishkova, S., and Dubrovsky, J. G. (2014). Arabidopsis homolog of trithorax1 (ATX1) is required for cell production, patterning, and morphogenesis in root development. *J. Exp. Bot.* 65, 6373–6384. doi: 10.1093/jxb/eru355

Nedelcu, A. M. (2019). Independent evolution of complex development in animals and plants: deep homology and lateral gene transfer. *Dev. Genes Evol.* 229, 25–34. doi: 10.1007/s00427-019-00626-8

Org, T., Chignola, F., Hetenyi, C., Gaetani, M., Rebane, A., Liiv, I., et al. (2008). The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. *EMBO Rep.* 9, 370–376. doi: 10.1038/sj.embor.2008.11

Ornelas-Ayala, D., Vega-Leon, R., Petrone-Mendoza, E., Garay-Arroyo, A., Garcia-Ponce, B., Alvarez-Buylla, E. R., et al. (2020). ULTRAPETALA1 maintains *Arabidopsis* root stem cell niche independently of ARABIDOPSIS TRITHORAX1. *New Phytol.* 225, 1261–1272. doi: 10.1111/nph.16213

Perez-Rodriguez, P., Riano-Pachon, D. M., Correa, L. G., Rensing, S. A., Kersten, B., and Mueller-Roeber, B. (2010). PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 38, D822–D827. doi: 10.1093/nar/gkp805

Pires, H. R., Monfared, M. M., Shemyakina, E. A., and Fletcher, J. C. (2014). ULTRAPETALA trxG genes interact with KANADI transcription factor genes to regulate Arabidopsis gynoecium patterning. *Plant Cell* 26, 4345–4361. doi: 10.1105/tpc.114.131250

Pu, L., Liu, M. S., Kim, S. Y., Chen, L. F., Fletcher, J. C., and Sung, Z. R. (2013). EMBRYONIC FLOWER1 and ULTRAPETALA1 act antagonistically on *Arabidopsis* development and stress response. *Plant Physiol.* 162, 812–830. doi: 10.1104/pp.112.213223

Roy, D., Chakrabarty, J., Mallik, R., and Chaudhuri, S. (2019). Rice Trithorax factor ULTRAPETALA 1 (OsULT1) specifically binds to "GAGAG" sequence motif present in Polycomb response elements. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862, 582–597. doi: 10.1016/j.bbagrm.2019.02.001

Saare, M., Rebane, A., Rajashekar, B., Vilo, J., and Peterson, P. (2012). Autoimmune regulator is acetylated by transcription coactivator CBP/p300. *Exp. Cell Res.* 318, 1767–1778. doi: 10.1016/j.yexcr.2012.04.013

Sanchez, M. P., Aceves-Garcia, P., Petrone, E., Steckenborn, S., Vega-Leon, R., Alvarez-Buylla, E. R., et al. (2015). The impact of Polycomb group (PcG) and Trithorax group (TrxG) epigenetic factors in plant plasticity. *New Phytol.* 208, 684–694. doi: 10.1111/nph.13486

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735–745. doi: 10.1016/j.cell.2007.02.009

Schuettengruber, B., Martinez, A. M., Iovino, N., and Cavalli, G. (2011). Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* 12, 799–814. doi: 10.1038/nrm3230

Shu, J., Chen, C., Thapa, R. K., Bian, S., Nguyen, V., Yu, K., et al. (2019). Genome-wide occupancy of histone H3K27 methyltransferases CURLY LEAF and SWINGER in *Arabidopsis* seedlings. *Plant Direct* 3:e00100. doi: 10.1002/pld3.100

Torok, M., and Etkin, L. D. (2001). Two B or not two B? Overview of the rapidly expanding B-box family of proteins. *Differentiation* 67, 63–71. doi: 10.1046/j.1432-0436.2001.067003063.x

Tyler, L., Miller, M. J., and Fletcher, J. C. (2019). The Trithorax group factor ULTRAPETALA1 regulates developmental as well as biotic and abiotic stress response genes in *Arabidopsis*. *G3* 9, 4029–4043. doi: 10.1534/g3.119.400559

Veraksa, A., Kennison, J., and McGinnis, W. (2002). DEAF-1 function is essential for the early embryonic development of *Drosophila*. *Genesis* 33, 67–76. doi: 10.1002/gene.10090

Wang, Y., Gu, X., Yuan, W., Schmitz, R. J., and He, Y. (2014). Photoperiodic control of the floral transition through a distinct polycomb repressive complex. *Dev. Cell* 28, 727–736. doi: 10.1016/j.devcel.2014.01.029

Whittaker, C., and Dean, C. (2017). The FLC locus: a platform for discoveries in epigenetics and adaptation. *Annu. Rev. Cell Dev. Biol.* 33, 555–575. doi: 10.1146/annurev-cellbio-100616-060546

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nat. Genet.* 49, 1546–1552. doi: 10.1038/ng.3937

Xie, Y., Straub, D., Eguen, T., Brandt, R., Stahl, M., Martinez-Garcia, J. F., et al. (2015). Meta-analysis of *Arabidopsis* KANADI1 direct target genes identifies a basic growth-promoting module acting upstream of hormonal signaling pathways. *Plant Physiol.* 169, 1240–1253. doi: 10.1104/pp.15.00764

Xu, F., Kuo, T., Rosli, Y., Liu, M. S., Wu, L., Chen, L. O., et al. (2018). Trithorax group proteins act together with a polycomb group protein to maintain chromatin integrity for epigenetic silencing during seed germination in *Arabidopsis*. *Mol. Plant* 11, 659–677. doi: 10.1016/j.molp.2018.01.010

Xu, L., and Shen, W. H. (2008). Polycomb silencing of KNOX genes confines shoot stem cell niches in *Arabidopsis*. *Curr. Biol.* 18, 1966–1971. doi: 10.1016/j.cub.2008.11.019

Yao, X., Feng, H., Yu, Y., Dong, A., and Shen, W. H. (2013). SDG2-mediated H3K4 methylation is required for proper *Arabidopsis* root growth and development. *PLoS One* 8:e56537. doi: 10.1371/journal.pone.0056537

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 in Plant Science

ORIGINAL RESEARCH
published: 09 April 2021
doi: 10.3389/fpls.2021.649634

Cotranscriptional and Posttranscriptional Features of the Transcriptome in Soybean Shoot Apex and Leaf

Jiafu Zhu[1], Han Zhao[2], Fanjiang Kong[1], Baohui Liu[1], Min Liu[1] and Zhicheng Dong[1]**

[1] *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] *Provincial Key Laboratory of Agrobiology, Institute of Crop Germplasm and Biotechnology, Jiangsu Academy of Agricultural Sciences, Nanjing, China*

Transcription is the first step of central dogma, in which the genetic information stored in DNA is copied into RNA. In addition to mature RNA sequencing (RNA-seq), high-throughput nascent RNA assays have been established and applied to provide detailed transcriptional information. Here, we present the profiling of nascent RNA from trifoliate leaves and shoot apices of soybean. In combination with nascent RNA (chromatin-bound RNA, CB RNA) and RNA-seq, we found that introns were largely spliced cotranscriptionally. Although alternative splicing (AS) was mainly determined at nascent RNA biogenesis, differential AS between the leaf and shoot apex at the mature RNA level did not correlate well with cotranscriptional differential AS. Overall, RNA abundance was moderately correlated between nascent RNA and mature RNA within each tissue, but the fold changes between the leaf and shoot apex were highly correlated. Thousands of novel transcripts (mainly non-coding RNA) were detected by CB RNA-seq, including the overlap of natural antisense RNA with two important genes controlling soybean reproductive development, *FT2a* and *Dt1*. Taken together, we demonstrated the adoption of CB RNA-seq in soybean, which may shed light on gene expression regulation of important agronomic traits in leguminous crops.

Keywords: soybean, chromatin-bound RNA, co-transcriptional splicing, non-coding RNA, nascent RNA

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Rui Xia,
South China Agricultural University, China
Yalong Guo,
Institute of Botany, Chinese Academy of Sciences, China

****Correspondence:***
Min Liu
minl@gzhu.edu.cn
Zhicheng Dong
zc_dong@gzhu.edu.cn

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

Received: *05 January 2021*
Accepted: *02 March 2021*
Published: *09 April 2021*

Citation:
Zhu J, Zhao H, Kong F, Liu B, Liu M and Dong Z (2021) Cotranscriptional and Posttranscriptional Features of the Transcriptome in Soybean Shoot Apex and Leaf. Front. Plant Sci. 12:649634. doi: 10.3389/fpls.2021.649634

INTRODUCTION

Transcription, the first step of gene expression, is accomplished by the multisubunit protein complex RNA polymerase. In eukaryotic cells, RNA polymerase II (RNA Pol II) is involved in protein-coding gene transcription and some non-coding gene transcription. Before maturation, messenger RNA precursors (pre-mRNAs) are subjected to multiple processing steps, including 5′ capping, splicing of introns, 3′ cleavage and polyadenylation, and editing (Bentley, 2014). These steps are known as posttranscriptional processing. However, increasing evidence suggests that most processes are cotranscriptional. For example, introns can be either co- or posttranscriptionally spliced, which is supported by the splicing loops of nascent RNA observed by electron microscopy in *Drosophila melanogaster* and *Chironomus tentans* (Beyer and Osheim, 1988; Baurén and Wieslander, 1994). In addition, high-throughput sequencing of nascent RNA revealed genome-wide cotranscriptional splicing (Khodor et al., 2011; Nojima et al., 2015; Drexler et al., 2020). Studies

from budding yeast, flies, and mammals indicated that cotranscriptional splicing frequencies are similarly high, ranging from 75 to 85% (Neugebauer, 2019).

Since Core et al. (2008) published a method wherein the nuclei run on RNA were affinity purified followed by high-throughput sequencing, nascent RNA sequencing (RNA-seq) technologies have significantly improved our ability to analyze transcription at each step across the genome. Rather than steady-state mRNA, nascent RNA-seq detects pre-mRNAs, divergent transcripts, enhancer-derived RNA (eRNA), etc., which are usually unstable and not polyadenylated. Recently, we and another laboratory have reported cotranscriptional splicing in the model plant *Arabidopsis* using genome-wide nascent RNA-seq approaches, plant native elongating transcript sequencing (pNET-seq), and plaNET-seq (Zhu et al., 2018; Kindgren et al., 2020). pNET-seq and plaNET-seq detect nascent RNA through enrichment of transcriptionally engaged RNA Pol II complexes, and splicing intermediates can also be observed when some spliceosomes are copurified with Pol II complexes (Zhu et al., 2018). Moreover, three recent publications directly sequenced the chromatin-bound RNA (CB RNA) of *Arabidopsis* and found genome-wide cotranscriptional splicing (Jia et al., 2020; Li et al., 2020; Zhu et al., 2020). However, the *Arabidopsis* genome is the first plant genome to be sequenced and is compact (140 million base/haploid genome), with an average gene length of 2,000 bp and an average intron length of 180 bp (Arabidopsis Genome Initiative, 2000). While harboring thousands to tens of thousands of genes, plant genome size ranges from approximately 0.1 to 100 gigabases (Pellicer and Leitch, 2020). Therefore, knowledge of transcription obtained from *Arabidopsis* may not be applicable to other plant genomes, especially some complicated crop genomes.

As one of the most important crops, soybean provides protein and oil for humans and livestock. During the past decades, great progress has been made in soybean genome research (Shen et al., 2018; Xie et al., 2019; Liu Y. et al., 2020). Furthermore, many important genes involved in agronomic traits have been characterized *via* genetic, cellular biology, and biochemical approaches (Kasai et al., 2007; Lu et al., 2017, 2020). For example, *Dt1*, which controls soybean growth habits, has been cloned as a *TFL1* homolog encoding a 173-amino-acid peptide (Liu et al., 2010; Tian et al., 2010). *FT2a* and *FT5a*, two distant homologous genes of *Dt1* within the same family, have been shown to play a conserved role in controlling flowering time (Kong et al., 2010; Takeshima et al., 2019; Wu et al., 2017).

Soybean [*Glycine max* (L.) Merr.] is a paleopolyploid derived from two whole genome duplication events approximately 59 and 13 million years ago. It has a relatively complicated and large genome, with a size of approximately 1.1 gigabases (Schmutz et al., 2010). The average gene length is approximately 4,000 bp, and the average intron length is approximately 539 bp in soybean (Shen et al., 2014), which are longer than those in *Arabidopsis*. Despite the considerable transcriptomic analyses of various soybean tissues using mature RNA-seq (Libault et al., 2010; Severin et al., 2010; Shen et al., 2014; Wang et al., 2014; Gazara et al., 2019), genome-wide analysis of nascent RNA from soybean has not yet been reported. In addition to capturing cotranscriptional features, nascent RNA is very sensitive to the detection of unstable regulatory RNAs, such as long non-coding RNAs (ncRNAs). Therefore, the investigation on nascent RNA in soybean would provide a comprehensive description of cotranscriptional characteristics in leguminous crops. Here, we report for the first time the analysis of nascent RNA from the shoot apex and leaf tissues of the soybean cv. Williams 82.

RESULTS

Nascent RNA Profiling of Soybean by CB RNA-Seq

The spatial and temporal expression of genes in the shoot apex largely determines the architecture of crop plants, including the numbers of branches, flowers, and nodes, which finally affect the yield per plant. Specifically, mRNA of *Dt1* was detected in the shoot apex at 15 days after emergence under a long-day condition (Liu et al., 2010); therefore, we set to investigate the transcriptome of the shoot apex from 10- to 15-day-old plants (**Figure 1A**, see section "Materials and Methods"). To gain insights of the shoot apex-specific gene, we chose the first trifoliolate leaves from 15-day-old plants as control. For nascent RNA, CB RNA was isolated, and the rRNA and polyA RNA it contained were depleted prior to library construction and high-throughput sequencing as described by Zhu et al. (2020). To further reveal cotranscriptional and posttranscriptional processes, we also conducted parallel mature polyA RNA-seq by enriching polyA RNA from total RNA, and these RNAs were constructed into libraries. Three biological replicates were sequenced and analyzed for each tissue. Principal component analysis (PCA) and Pearson correlation analysis of gene expression indicated high reproducibility of biological replication (**Figure 1B** and **Supplementary Figure 1**). In addition, the first two components of PCA explained more than 90% of the variation, indicating that the tissue difference (apex vs. leaf, 61.81% of variance) and methodological difference (CB RNA-seq vs. polyA RNA-seq, 28.46% of variance) were the dominant factors for intersample differentiation (**Figure 1B**).

As expected, the read distribution of nascent RNA shows two characteristics compared with that of polyA RNA. First, CB RNA-seq detected more intron signals than polyA RNA-seq because more unspliced reads were sequenced at the nascent RNA level. Approximately 25% of unique mapped reads were located in the intron region with CB RNA-seq, while less than 4% of unique mapped reads were located in the intron region with polyA RNA-seq (**Supplementary Figure 2**). In addition, the read density ratio of introns to exons in CB RNA was significantly higher than that in polyA RNA (**Figure 1C**). Second, the read density on the gene decreased gradually from the 5′ end to the 3′ end, while there was no such phenomenon in polyA RNA (**Figures 1D,E**). For example, the read signal of the gene *Glyma.02G231800* declined from 5′ to 3′ in CB RNA-seq but not in polyA RNA-seq. Furthermore, an intron signal was evident in CB RNA but absent from polyA RNA (**Figure 1D**). These characteristics were consistent with the results from previous studies and confirmed

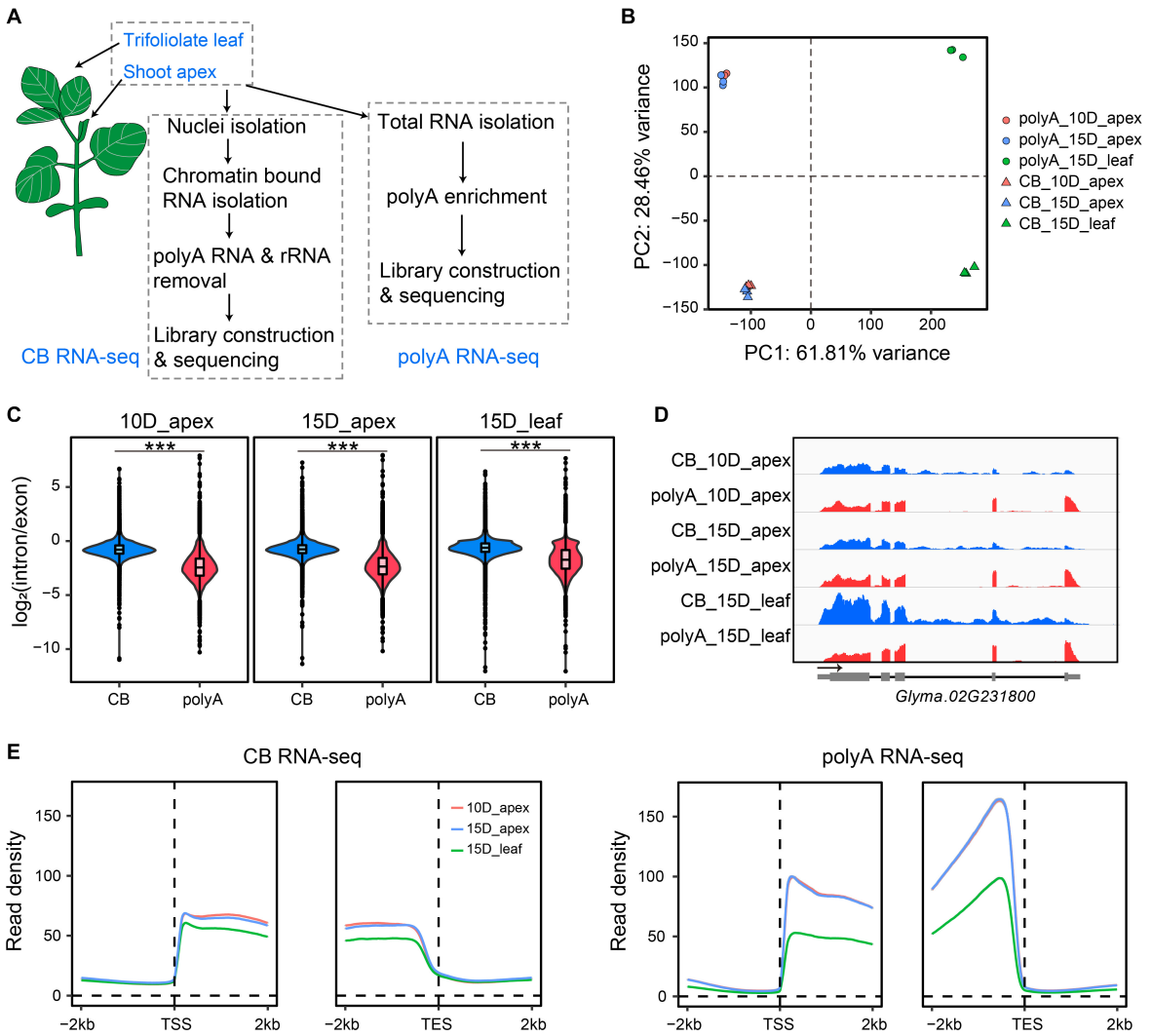

FIGURE 1 | Overview of the experimental design and features of nascent RNA and mRNA. **(A)** Scheme of chromatin-bound RNA sequencing (CB RNA-seq) and polyA RNA-seq. **(B)** Principal component analysis (PCA) of gene expression of biological triplicates from CB RNA-seq and polyA RNA-seq. The triangles and dots represent CB RNA-seq and polyA RNA-seq, respectively. Red, 10-day apex; blue, 15-day apex; green, 15-day leaf. **(C)** Comparison of the gene intron/exon ratio between CB RNA-seq and polyA RNA-seq (left, 10-day apex; middle, 15-day apex; right, 15-day leaf). ***$p < 0.001$, Wilcoxon test. **(D)** Screenshot of IGV showing the read distribution of CB RNA-seq and polyA RNA-seq on the *Glyma.02G231800* gene. Blue, CB RNA-seq; red, polyA RNA-seq. **(E)** Profiles of read density of CB RNA-seq (left) and polyA RNA-seq (right) for the 2-kb up- and downstream transcription start site (TSS) and transcription end site (TES). Lines represent the mean value of read density. Ten-day apex, 15-day apex, and 15-day leaf samples are indicated in red, blue, and green, respectively.

that the CB RNAs obtained here were bona fide transcriptional processing nascent RNAs (Li et al., 2020; Zhu et al., 2020).

Multiple Factors Regulate Cotranscriptional Splicing Efficiency

Cotranscriptional splicing has been widely found in eukaryotic cells. We wondered whether splicing coupled with transcription is widespread in the soybean genome. The intron retention ratio is an indicator of intron splicing efficiency. Thus, we adopted an index for the percent of intron retention (PIR) to measure the extent of cotranscriptional splicing (Braunschweig et al., 2014). In short, the PIR of an intron was calculated as the ratio of unspliced exon–intron junction reads to the total junction reads (unspliced exons–introns and spliced exons–exons). Since each unspliced exon–intron read from one RNA molecule has the chance to be sequenced twice in high-throughput sequencing, the average count of exon–intron reads at the 5′ splice site (EI5) and of exon–intron reads at the 3′ splice site (EI3) was considered an intron's unspliced exon–intron read count (**Figure 2A**). Introns with lower PIR values are more efficient for splicing. Constitutive introns of active genes (TPM > 1) were calculated for PIR both in CB RNA and polyA RNA. As expected, the intron retention levels of CB RNA were significantly higher than those of polyA RNA, both in the apex and leaf (**Figure 2B**). Most introns in polyA RNA have a very low PIR, usually smaller than 0.1. The

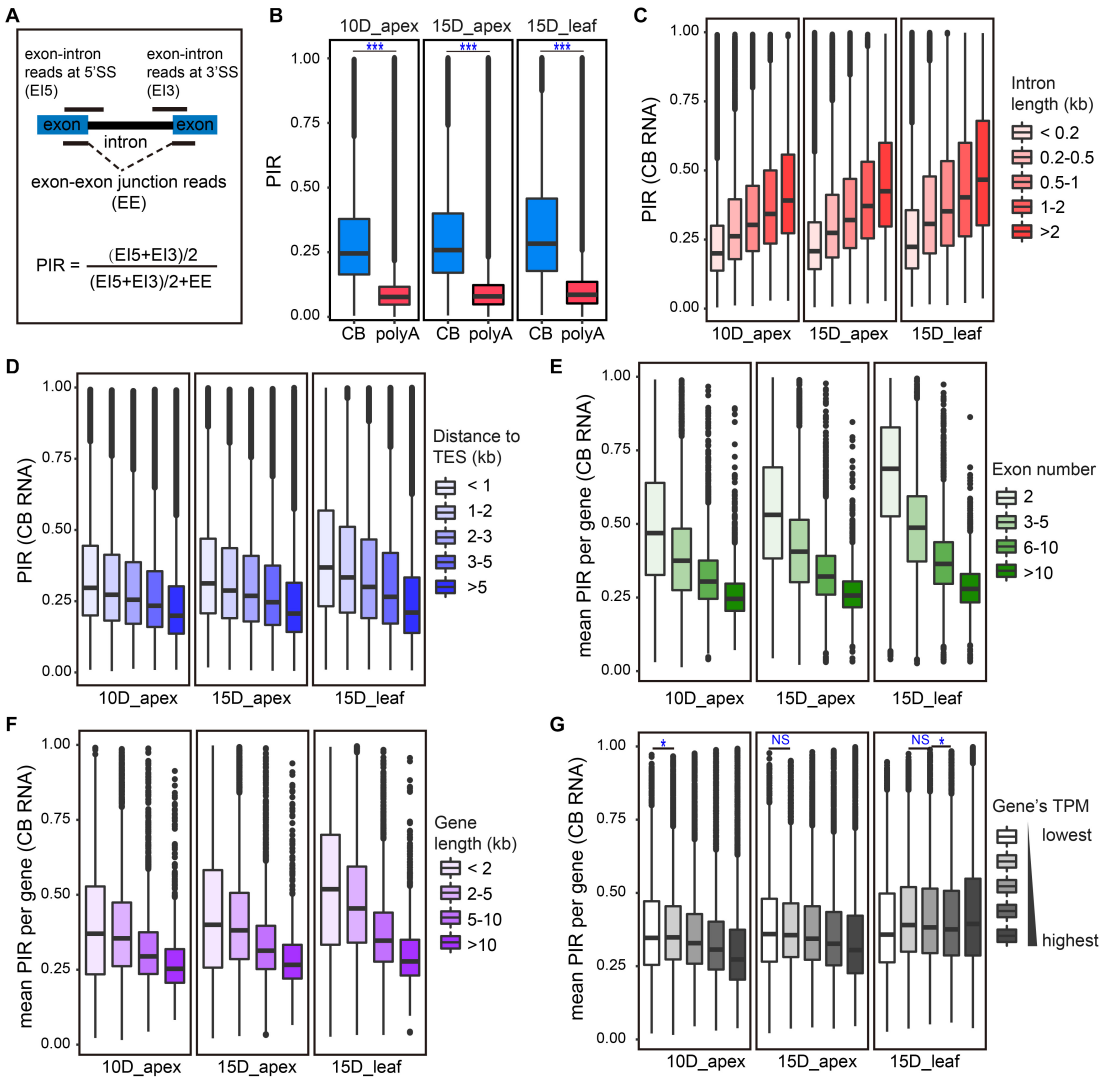

FIGURE 2 | CB RNA-seq detected cotranscriptional splicing processing. **(A)** Calculation of percent of intron retention (PIR). 5′SS, 5′ splice site; 3′SS, 3′ splice site. **(B)** Boxplots of the overall PIR of CB RNA and polyA RNA for 10D_apex (left), 15D_apex (middle), and 15D_leaf (right). ***$p < 0.001$, Wilcoxon test. **(C,D)** Boxplots of PIR levels of introns of different sizes **(C)** and distances from the transcription end site (TES) **(D)**. **(E–G)** Boxplots of the average PIR from genes with different exon numbers **(E)**, gene lengths **(F)**, and expression levels **(G)**. For **(C–G)**, the Wilcoxon test was used to test the difference in PIRs for adjacent groups. All tests were highly significant ($p < 0.001$) unless symbols were assigned (*$p < 0.05$; NS, $p > 0.05$).

median PIR was close to 0.25 (in the apex) or above 0.25 (in the leaf) in CB RNA. These results were similar to those of a previous study of *Arabidopsis* (Li et al., 2020). The PIR of most introns in CB RNA was lower than 0.5 (PIR = 1 means completely unspliced), indicating the existence of genome-wide cotranscriptional splicing in soybean.

Although most introns undergo cotranscriptional splicing, the extent of intron retention is highly variable. Studies in *Drosophila* and *Arabidopsis* have indicated that multiple factors, such as intron characteristics, gene expression level, and number of introns, are related to cotranscriptional splicing efficiency (Khodor et al., 2011; Li et al., 2020; Zhu et al., 2020). To examine how these factors affect the splicing efficiency in soybean, we first divided introns into five groups by length and found that intron retention became more prominent as the intron length increased (**Figure 2C**).

In addition to intron length, the intron position is also supposed to influence splicing efficiency. According to the "first come, first served" model, there may be more splicing chances for introns transcribed first (Aebi et al., 1986). Based on the distance to transcription end sites (TES), introns were divided into five groups, and the PIR was compared among groups. Introns more distant from TES are transcribed early and thus are more likely to be spliced first. As expected, the PIR index gradually declined as the intron distance to TES decreased (**Figure 2D**).

In addition, the cotranscriptional splicing efficiency was positively correlated with exon number (**Figure 2E**) and gene length (**Figure 2F**). These patterns were consistent between the

apex and leaf tissues. However, a weak positive correlation of cotranscriptional splicing and gene expression was detected in the apex instead of in the leaf (**Figure 2G**).

Cotranscriptional Splicing Efficiency Is Correlated With Certain Histone Modifications

Specific histone modifications have been shown to regulate cotranscriptional splicing by either directly recruiting spliceosomes or indirectly influencing transcriptional elongation (Luco et al., 2010; Hu et al., 2020). To test whether cotranscriptional splicing is associated with certain histone modifications in soybean, we used ChIP-seq data of several histone modifications (H3K27me3, H3K4me1, H3K4me3, H3K36me3, H3K56ac, and H2A.Z) in leaf tissue collected from a previous study (**Supplementary Table 1**; Lu et al., 2019). We then quantified the level of different histone modifications around introns in different groups based on the retention rates (**Figure 3**). PIR is positively correlated with the levels of H3K27me3, H3K4me3, H3K56ac, and H2A.Z-marked histone, which means that introns with higher cotranscriptional splicing efficiency have lower levels of those histone modifications. PIR is negatively correlated with the level of H3K4me1-marked histones. Notably, H3K27me3, H3K4me3, H3K56ac, H3K36me3, and H2A.Z showed a higher modification level at the upstream exon than at the downstream exon, while H3K4me1 showed a higher modification level at the downstream exon. It is most likely that these histone modifications, H3K27me3, H3K4me3, H3K56ac, H3K36me3, and H2A.Z, preferentially locate at the gene's 5′ end, except for H3K4me1 (**Supplementary Figure 3**).

Alternative Splicing Events Are Likely Determined Cotranscriptionally

In higher eukaryotes, alternative splicing (AS), as an important regulatory step of gene expression, plays a critical role in the development and stress response of organisms (Baralle and Giudice, 2017; Laloum et al., 2018). Previous studies in mammalian cells and *Arabidopsis* showed that AS events occur co- or post-transcriptionally (Jia et al., 2020). Thus, we wondered to what extent AS is determined cotranscriptionally. We adopted percent spliced-in (PSI) (Wang et al., 2008) to describe the relative abundance of splicing events. We focused on four AS events: alternative 3′ splice sites (A3SS), alternative 5′ splice sites (A5SS), exon skipping (ES), and retained introns (RI) (**Figure 4A**). The PSI values of AS events from CB RNA and polyA RNA were significantly correlated, suggesting that AS events are likely determined cotranscriptionally for all AS types (**Figure 4B**). This was true for both shoot apex and leaf tissues (**Figure 4** and **Supplementary Figure 4**). However, the overall PSI value was higher in CB RNA (**Figure 4B**, insets). For AS events with a higher PSI in CB RNA than in polyA RNA, there are two possible explanations. First, some highly abundant transcripts in CB RNA with AS events may likely be rapidly degraded. For example, coupling of AS and nonsense-mediated mRNA decay (NMD) has been reported to fine-tune gene expression (McGlincy and Smith, 2008). Second, posttranscriptional splicing may lead to a higher PSI in CB RNA, especially for RI events.

Differential Alternative Splicing Between Leaf and Shoot Apex Tissues Is Not Determined Merely by Cotranscriptional Splicing

Given that most AS events are determined cotranscriptionally, we then asked whether differences in AS between the shoot apex and leaf tissues detected by CB RNA-seq and polyA RNA-seq are consistent. Thus, we compared the AS difference of both CB RNA and polyA RNA between the 15-day apex and leaf tissues. Differential splicing events were analyzed by the program SUPPA2 (Trincado et al., 2018). A splicing event was considered differential when the absolute value of the PSI difference (ΔPSI) between tissues >0.1 and the p-value < 0.05. A small number of the different splicing events between the leaf and shoot apex tissues were detected by both CB RNA and polyA RNA (**Figure 5A**). ΔPSI_{mRNA} and ΔPSI_{CB} were barely correlated (Spearman correlation ranged from 0.22 to 0.35) (**Figure 5B**). Furthermore, genes with different splicing events detected by CB RNA and polyA RNA were not concordant (**Supplementary Figure 5A**). Although overall AS events are highly correlated at the cotranscriptional level and posttranscriptional level within the same tissue, tissue-specific mRNA processing, such as degradation and posttranscriptional splicing, may result in the differential AS events that are detected by polyA RNA but not by CB RNA. For those differential AS events detected by CB RNA but not by polyA RNA, it was probably caused by the differentially cotranscriptional splicing efficiency between the shoot apex and leaf tissues and further corrected at the posttranscriptional splicing step, exemplified by the first intron of *Glyma.07G206100* (**Supplementary Figure 6**).

Genes associated with intertissue differential splicing events detected by CB RNA and polyA RNA were also different (**Supplementary Figure 5A**). To explore the biological function of genes with different AS events, we conducted Gene Ontology (GO) enrichment analysis. Interestingly, genes with different splicing events between the 15-day apex and leaf tissues were significantly enriched in mRNA splicing and RNA processing, which somehow explains the differential splicing efficiency between the shoot apex and leaf tissues (**Supplementary Figure 5B**).

The Level of Steady-State mRNA Is Moderately Correlated With the Biogenesis of Nascent RNA

Chromatin-bound RNA-seq is applied to detect transcribed RNAs, which are subject to multiple steps of mRNA processing, including cotranscriptional and posttranscriptional processes prior to maturation. Thus, there might be discordance in the abundance at the nascent RNA and mRNA levels. To test this hypothesis, we compared the TPM values of nascent RNA and mature RNA. Overall, the levels of nascent RNA and mature RNA were moderately correlated (Spearman correlation = 0.71–0.73) (**Figure 6A** and **Supplementary Figures 7A–C**). There are two

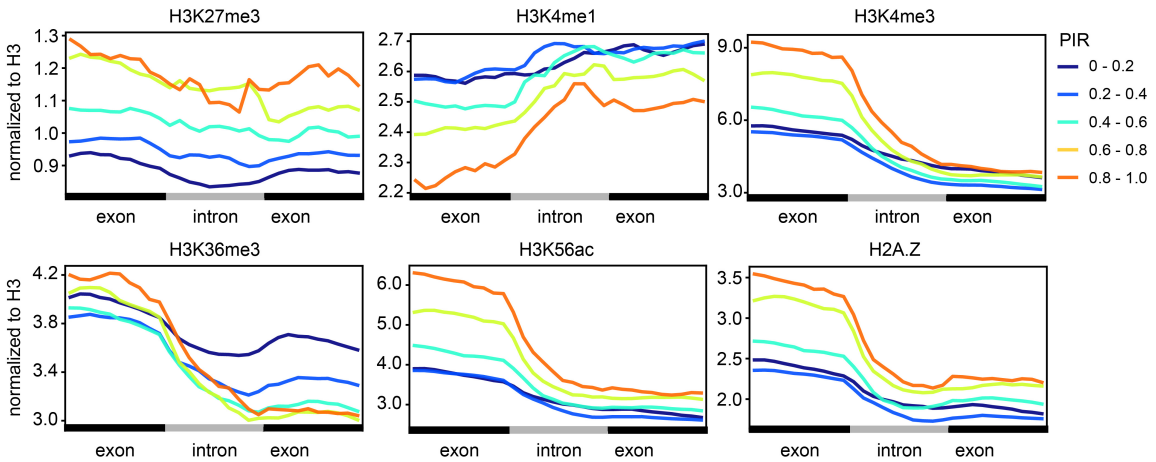

FIGURE 3 | Cotranscriptional splicing efficiency is correlated with certain histone modifications. Levels (*y*-axis) of different histone modifications (H3K27me3, H3K4me1, H3K4me3, H3K36me3, H3K56ac, and H2A.Z) along introns and the flanking exons (*x*-axis). Lines with different colors indicate intron groups divided according to the PIR. The ChIP-seq data of different histone modifications in leaf tissue were adopted from a previous study (**Supplementary Table 1**; Lu et al., 2019).

types of discordant genes. One is a gene that is highly transcribed with a low level of mature RNA, which might result from a high turnover of mRNA and is designated unstable RNA. The other is a gene with relatively low transcription activity but a high level of mature RNA, which might be due to the high RNA stability and is called stable RNA.

To select unstable and stable RNA transcripts, we first established a linear regression model of the log2 values of TPM genes obtained with CB RNA-seq and polyA RNA-seq. Then, the predicted TPM values of genes in polyA RNA were calculated based on the linear regression model. If the actual TPM of a gene was threefold higher (or lower) than the predicted TPM, the gene was considered to be stable (or unstable) (**Figure 6A** and **Supplementary Figures 7B,C**). To investigate whether the stability of RNA is associated with specific biological functions, we performed GO enrichment analysis. For unstable RNAs, defense response, protein phosphorylation, and signal transduction were the most enriched terms. Stable RNAs were mainly associated with translation, photorespiration, ribosome biogenesis, and glycolytic processes (**Figure 6B** and **Supplementary Figures 7C,E**).

Differentially Expressed Genes Are Consistent at the Nascent and Mature RNA Levels

We then identified differentially expressed genes (DEGs) between 15-day apex and 15-day leaf tissues at both nascent and mature RNA levels. More than 10,000 genes were expressed more in the apex than in the leaf, and vice versa (**Supplementary Figure 8A** and **Supplementary Table 2**). Most of these DEGs detected by CB RNA-seq and polyA RNA-seq overlapped (**Figure 6C** and **Supplementary Figure 8B**). Furthermore, fold changes at the CB RNA level and polyA RNA level were highly correlated (Spearman correlation = 0.93) (**Figure 6D**).

Gene Ontology enrichment analysis was performed to determine the biological functions of the DEGs. Genes with higher expression in the apex were mainly associated with RNA methylation, histone methylation, translation, DNA replication, and meristem initiation and maintenance. Genes with higher expression levels in the leaves were mainly related to photosynthesis and plastid organization (**Supplementary Figure 8C**).

In addition, only a small number of genes were called DEGs between the 15-day apex and 10-day apex (**Supplementary Figure 9A**), and they had concordant changes at the nascent RNA and mRNA levels (**Supplementary Figure 9B**). GO enrichment indicated that genes highly expressed in the 10-day apex were involved in the response to stress, circadian rhythm, etc., and genes highly expressed in the 15-day apex were involved in long-day photoperiodism flowering, response to hormones, and circadian rhythm (**Supplementary Figure 9C**).

More Non-coding RNAs Were Identified by CB RNA-Seq Than PolyA RNA-Seq

Considering that unstable transcripts are readily detected at the nascent RNA level, we calculated the expression level of ncRNA as defined in a previous study (Lin et al., 2020). As expected, more active ncRNA genes were detected by CB RNA-seq than polyA RNA-seq (**Figure 7A**). Furthermore, we determined the antisense transcription of annotated mRNAs by counting reads mapped to the opposite strand, and there were more active antisense transcriptional signals at the nascent RNA level (**Figure 7B**, left). These results indicate that some non-coding transcripts were unstable or not polyadenylated. For example, a transcript encoded from the antisense strand of *FT2a*, the essential gene involved in flowering timing, was identified in 15-day leaves by CB RNA-seq. *Dt1*, the key gene controlling growth habit, overlapped with another strong antisense transcript at the nascent RNA level in the apex (**Figure 7B**, right).

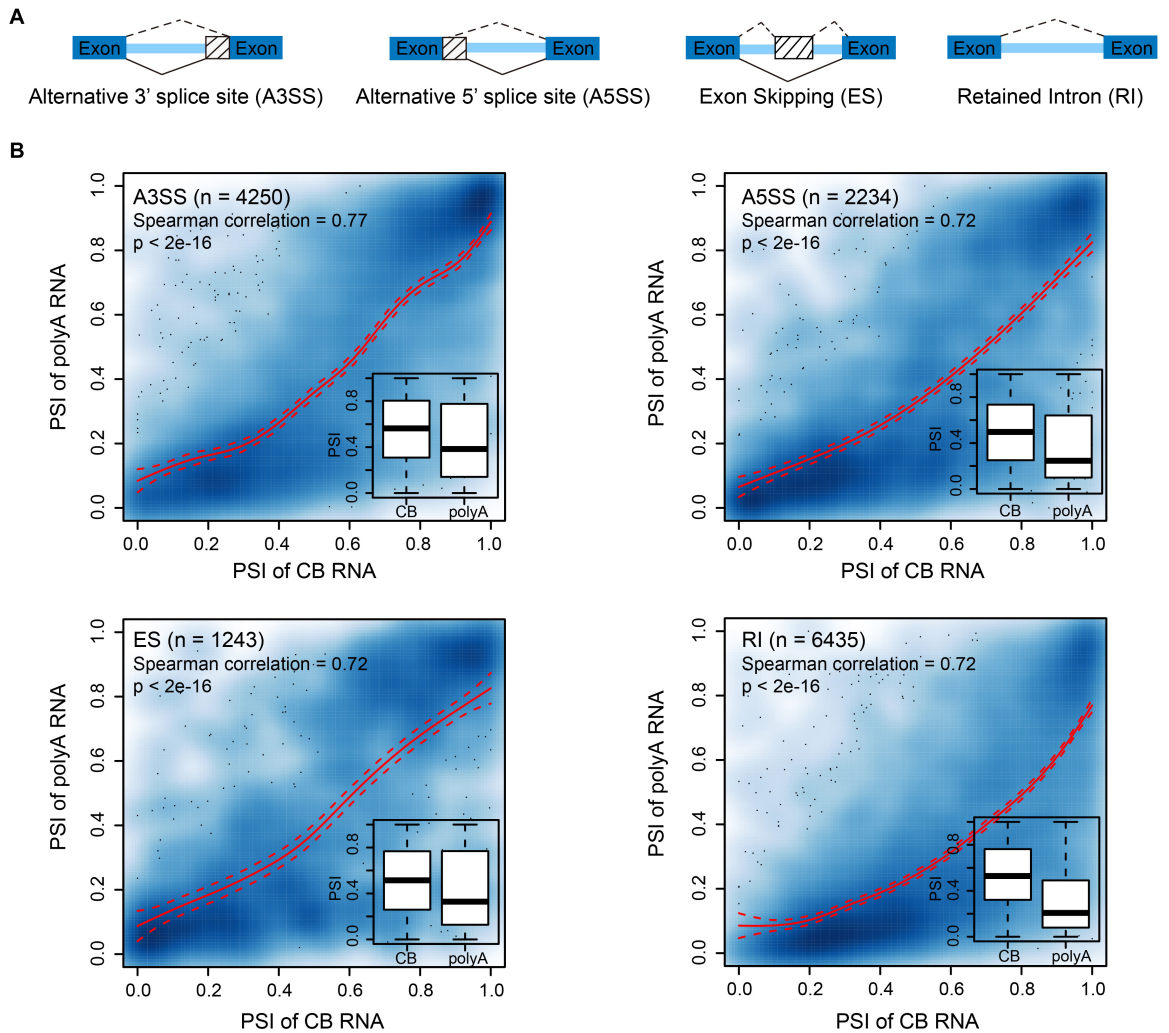

FIGURE 4 | Alternative splicing events are likely determined cotranscriptionally. **(A)** Diagram showing the different alternative splicing events analyzed. **(B)** Scatter plots showing the correlation between the percent spliced-in (PSI) values of CB RNA and polyA RNA of different AS events in 15-day apex tissues. Smooth spline curves were fitted (solid red lines), and 95% confidence intervals were plotted (dashed red lines). Spearman's correlation coefficients are presented in the plots. Insets show boxplots of PSIs for AS events at the CB RNA and polyA RNA levels.

To identify novel transcripts, we assembled transcripts from nascent RNA and polyA RNA of each tissue separately. Then, all transcripts were merged and compared based on reference annotations (see section "Materials and Methods"). Only intergenic transcripts were included for further analysis. In total, there were 5,927 and 1,515 active intergenic transcripts from CB RNA-seq and polyA RNA-seq, respectively, with 1,326 transcripts overlapping (**Figure 7C**, upper panel; **Supplementary Table 3**). These transcripts were encoded from 4,835 loci, of which 1,142 were shared by CB RNA and polyA RNA (**Figure 7C**, bottom panel).

We then applied two tools, CNCI and FEELnc, to evaluate the protein-coding potential of these new transcripts. In total, 4,001 and 974 active new transcripts of CB RNA and polyA RNA were considered non-coding transcripts by both methods, respectively (**Figure 7D**), and more ncRNAs were observed in the leaves at the nascent RNA level (**Figure 7E**).

Non-coding RNA detected only at the nascent RNA level might be unstable or unpolyadenylated. ncRNAs detected only at the polyA RNA level might be very stable and accumulate by slow transcription. Different types of ncRNAs may be regulated differently at the transcriptional level. To gain insight into the effects of histone modifications on ncRNA expression, we compared the metaprofiles of histone modifications for three groups of ncRNAs from the leaf tissue (group I: only detected by CB RNA; group II: detected by both; group III: only detected by polyA RNA) (**Figure 7F**). Group II and III ncRNA genes were associated with H3K56ac, H3K4me3, and histone variant H2A.Z (**Figure 7G**).

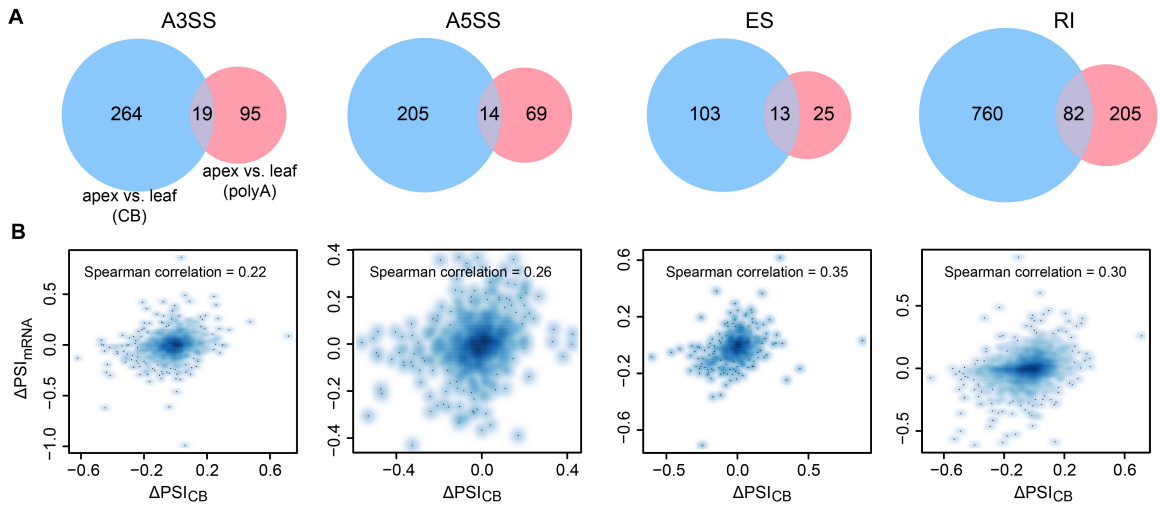

FIGURE 5 | Differential AS events detected by CB RNA and polyA RNA between the 15-day apex and leaf tissues. **(A)** Venn diagram illustrating the different AS events analyzed using CB RNA and polyA RNA. **(B)** Scatter plots show the correlation of ΔPSI of CB RNA (ΔPSI_{CB}) and polyA RNA (ΔPSI_{mRNA}) for AS events. Spearman correlations are indicated.

DISCUSSION

Although nascent RNA-seq has been extensively used to detect cotranscriptional regulation in yeast, fly, and mammalian cells, its application in plants is still lagging behind. Recently, several methods have been developed to detect nascent RNA and reveal plant-specific transcriptional features (Hetzel et al., 2016; Zhu et al., 2018). However, with the exception of one maize publication using GRO-seq (Erhard et al., 2015), all studies have focused on the model plant *Arabidopsis*. Here, we describe the soybean transcriptome using CB RNA-seq. As expected, CB RNA isolation greatly enriched the nascent RNA by removing the abundant cytosolic mRNAs and nucleoplasmic RNAs. We demonstrated that CB RNA-seq successfully detected nascent RNA biogenesis and cotranscriptional processing of pre-mRNA from the leaves and growing apex tissues. This method can be applied to other tissues at various developmental stages and/or under different environmental conditions, which may further shed light on the transcriptional regulation of the soybean genome.

We found genome-wide cotranscriptional splicing in soybean. Cotranscriptional splicing efficiency is related to intron length, distance from TES, intron number, and gene length. These characteristics are similar to those previously observed in yeast, fly, mammalian, and *Arabidopsis* cells, indicating a conserved mechanism that controls cotranscriptional splicing in eukaryotic cells (Khodor et al., 2011; Kindgren et al., 2020; Li et al., 2020). Interestingly, we found that both active (H3K4me3 and H3K56ac) and inactive (H3K27me3) histone markers are negatively related to cotranscriptional splicing efficiency. The elongation rate of RNA Pol II can affect splicing efficiency by fine-tuning the timing of the spliceosome search for splice sites, as the spliceosome is physically recruited by the carboxyl terminal domain of the largest subunit of RNA Pol II (Nojima et al., 2018). The inverse correlation between elongation speed and splicing efficiency was proven in yeast *in vivo* (Carrillo Oesterreich et al., 2016; Aslanzadeh et al., 2018). Moreover, the RNA Pol II elongation rate is regulated by transcription elongation factors and chromatin structural barriers such as nucleosomes. Thus, factors that affect transcription elongation also affect splicing efficiency. Active histone markers are thought to be related to a higher transcription elongation rate. Therefore, it is reasonable that introns with higher H3K4me3 or H3K56ac contents are less efficiently spliced. In addition, the pattern described in this study and a previous study on *Arabidopsis* revealed that the retained introns are derived from genes with low H3K4me1 and high H3K27me3 signatures (Mahrez et al., 2016). However, further studies of mutants with impaired histone modification are needed to verify their function in cotranscriptional splicing. Actually, these effects are not unidirectional. Cotranscriptional splicing can in turn influence the elongation rate and establishment of histone modifications (Kim et al., 2011).

Alternative splicing is an important part of gene regulation. In our study, a highly correlated relative AS event (PSI) was observed between CB RNA and polyA RNA, suggesting that most AS events are determined cotranscriptionally. This agrees with a previous study in *Arabidopsis* (Zhu et al., 2020). However, when comparing intertissue AS events, differential AS events detected at the cotranscriptional and posttranscriptional levels only partially overlapped. Thus, differential AS events cannot be predicted at the nascent RNA level, indicating the complexity of AS regulation. These regulations may be attributed to different degradation rates and/or posttranscriptional splicing among various tissues.

Gene expression is regulated at multiple levels, including transcription, post-transcription, and translation. Steady-state mRNA is the output of transcriptional activity and RNA degradation. Thus, there might be some discordance in gene

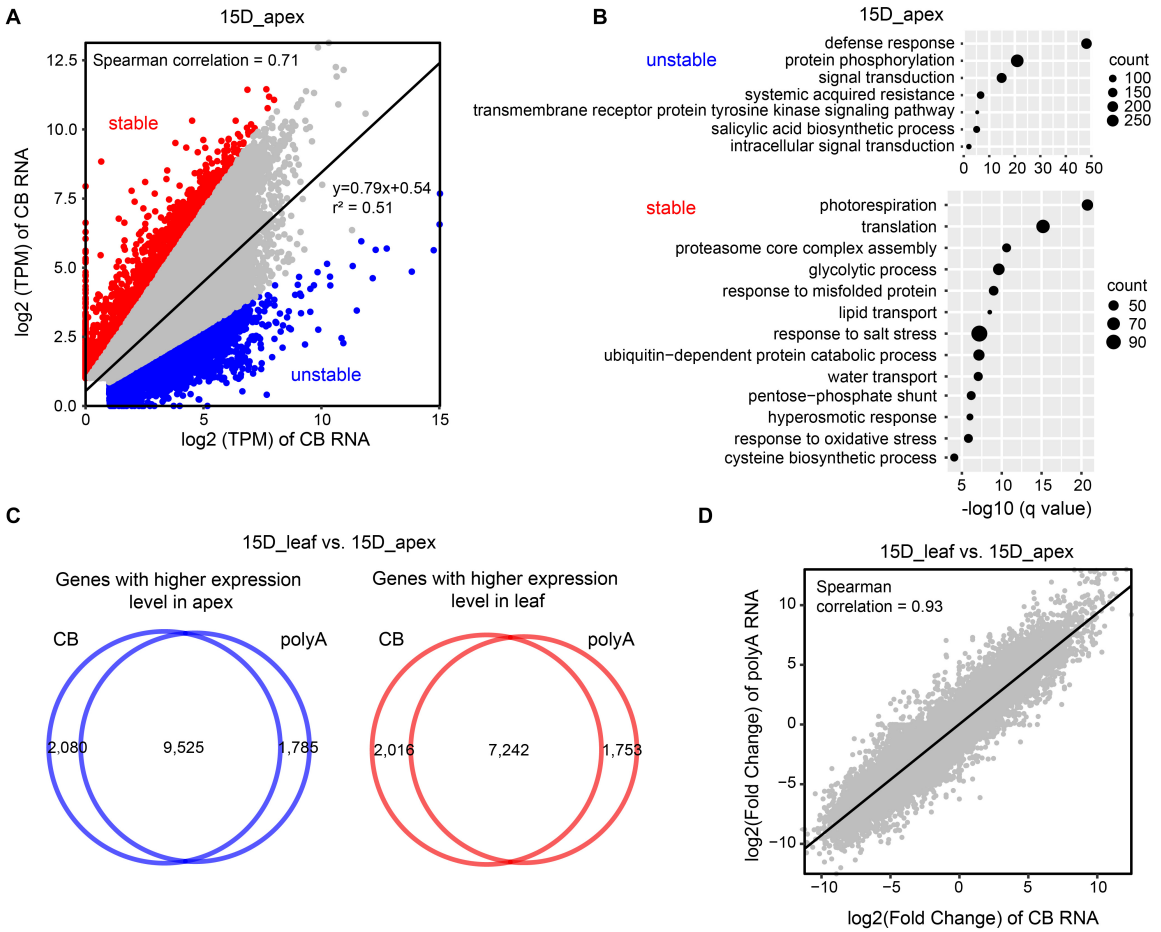

FIGURE 6 | Comparison of the expression levels of CB RNA and polyA RNA. **(A)** Scatter plot of gene TPM values (in log2) detected by CB RNA-seq and polyA RNA-seq in a 15-day apex. Spearman correlation and linear regression (solid line) equations are shown. Stable and unstable transcripts are represented by red and blue dots, respectively. **(B)** Gene Ontology (GO) enrichment of genes with unstable (upper) and stable RNA (bottom). **(C)** Venn diagrams show the number of genes with higher expression levels in 15-day apexes (left) and 15-day leaves (right) at the CB RNA and polyA RNA levels. These genes were detected by comparing DEGs between 15-day apex and 15-day leaf tissues (| fold change| $>$ 2, q-value $<$ 0.05). **(D)** Scatter plot of fold change of gene expression (15-day apex vs. 15-day leaf) detected by CB RNA-seq and polyA RNA-seq. The Spearman correlation value is presented.

activity detected by nascent RNA-seq and polyA RNA-seq. As expected, we found that gene activity at these two levels was moderately correlated. However, when comparing different tissues, the changes in gene activity at both levels were highly consistent, indicating that tissue-specific gene expression was mainly associated with transcription. The stability of RNA might contribute to the discordance in gene activity at the nascent and mature RNA levels. It is meaningful for stable mRNA genes to be involved in housekeeping biological processes. Moreover, under normal conditions, keeping regulatory genes at low mRNA levels and relatively high transcription by fast turnover of mRNA is an effective way to ensure rapid responses to potential stimuli. As we have previously reported in *Arabidopsis*, genes induced highly and quickly by short-term heat shock usually exhibit basic transcription under normal temperature (Liu M. et al., 2020).

Since some ncRNAs are unstable or unpolyadenylated, such as enhancer RNAs and antisense RNAs, more transcripts are expected to be detected by CB RNA-seq. However, this does not rule out the possibility that some transcripts detected only in CB RNA are not nascent RNA but rather chromatin-bound transcripts. To further elucidate the biological significance of these ncRNAs, approaches such as RNA interference and gene editing are needed. It will be interesting to apply CB RNA-seq to various tissues and build a transcriptional regulatory network at the nascent RNA level in the future.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Soybean Wm82 plants were grown under long light day conditions (16 h light, 8 h dark) with a constant 25°C temperature in a growth chamber. Shoot apexes from 10- to 15-day seedlings were collected in three biological replicates, with each replicate collected from approximately 20 plants. For the leaves, the first trifoliolate leaves of two 15-day-old plants were collected as one

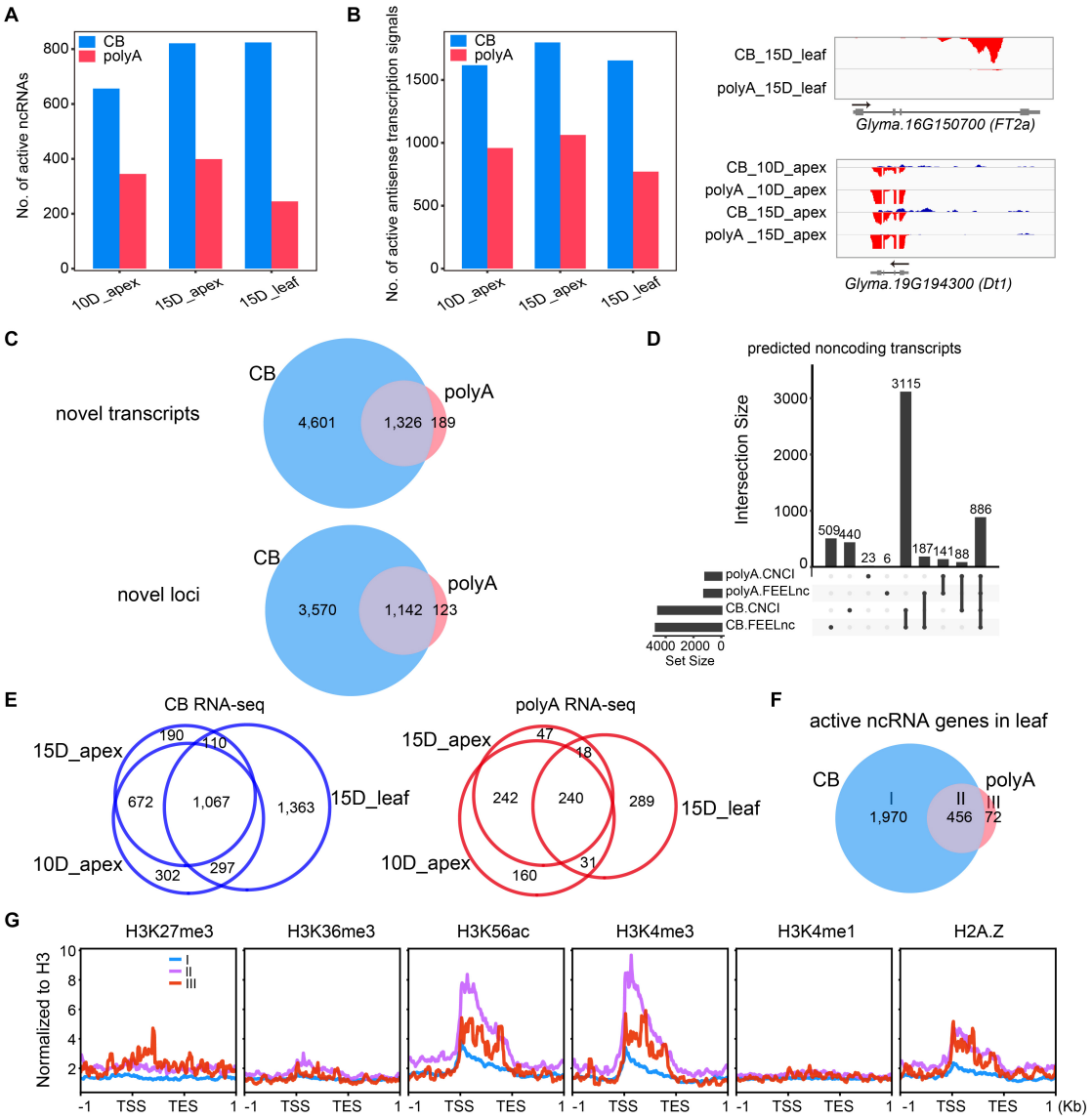

FIGURE 7 | CB RNA-seq detected more ncRNA transcripts than polyA RNA-seq. **(A)** Number of active ncRNAs detected by CB RNA-seq and polyA RNA-seq. **(B)** Number of active antisense transcription signals detected by CB RNA-seq and polyA RNA-seq (left). Two examples of ncRNAs are shown in the IGV screenshot (right). **(C)** Venn diagrams show the novel transcripts (upper) or the related loci (bottom) detected by CB RNA-seq and polyA RNA-seq. **(D)** UpSet plot shows the number of ncRNAs defined by CNCI and FEELnc in CB RNA and polyA RNA. **(E)** Venn diagrams show the overlapping ncRNA genes of the shoot apex and leaf tissues at the CB RNA (left) and polyA RNA levels (right). **(F)** Venn diagram showing three types of active ncRNAs in the leaves. Group I, only detected by CB RNA; group II, detected by both; group III, only detected by polyA RNA. **(G)** The average distribution of different histone modifications in each ncRNA group.

biological replicate. All samples were frozen with liquid nitrogen immediately after collection.

RNA Isolation, Transcriptome Library Preparation, and Sequencing

The chromatin RNA extraction protocol was modified from a previously published method (Zhu et al., 2020). Briefly, tissues were ground into a fine powder with liquid nitrogen and solubilized in cold nuclei isolation buffer (20 mM/KOH pH 7.4, 0.44 M sucrose, 1.25% Ficoll, 2.5% Dextran T40, 10 mM $MgCl_2$, 0.75% Triton X-100, 0.5 mM EDTA, 1 mM DTT, 8 mM β-mercaptoethanol, 1 μg/ml pepstatin A, 1 μg/ml aprotinin, and 1 mM PMSF). The crude nuclei were precipitated at 3,500 rpm and washed with resuspension buffer (50% glycerol, 25 mM Tris–HCl pH 7.5, 0.5 mM EDTA, 100 mM NaCl, 1 mM DTT, 0.4 U/μl RNase inhibitor, 1 μg/ml pepstatin A, 1 μg/ml aprotinin, and 8 mM β-mercaptoethanol) once, followed by washing buffer (25 mM Tris–HCl pH 7.5, 300 mM NaCl, 1 M urea, 0.5 mM EDTA, 1 mM DTT, 1% Tween-20, 0.4 U/μl RNase inhibitor, 1 μg/ml pepstatin A, 1 μg/ml aprotinin, and 8 mM

β-mercaptoethanol) twice. Chromatin RNA was extracted from washed nuclei using TRIzol reagent (Life Technologies).

After degrading genomic DNA by TURBO DNase (Life Technologies), CB RNA was subjected to rRNA depletion using a riboPOOL kit (siTOOLs Biotech, PanPlant-10 nmol) and polyA RNA removal by oligo(dT) beads (NEB, S1419). Poly(A) RNA was enriched from total RNA by oligo(dT) beads. Both CB RNA and polyA RNA were transformed into cDNA libraries using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7765) and sequenced on an Illumina NovaSeq platform.

CB RNA and mRNA Data Processing

Raw reads of CB RNA and polyA RNA were first evaluated by FastQC[1], and then Cutadapt was used to remove adapters and low-quality reads (Martin, 2011). Clean reads were subsequently aligned to the genome Wm82.a2.v1 by STAR (Dobin et al., 2013). Only uniquely mapped reads were retained for the following analysis. Read distribution on genomic features was evaluated by RSeQC with the subcommand "read_distribution.py" (Wang et al., 2012). To calculate the ratio of introns vs. exons of each gene, featureCounts was used to quantify the read counts on introns and exons separately (Liao et al., 2014). Read density was normalized by the length of introns and exons.

Calculating the Percent of Intron Retention

The proportion of intron-retained reads across an intron is usually used to evaluate the splicing efficiency of the intron. To quantitatively evaluate the genome-wide cotranscriptional splicing efficiency in soybean, we calculated the PIR value for constitutive introns as described previously (Braunschweig et al., 2014). Briefly, three types of reads on an intron were counted: (1) exon–intron junction reads across the 5′SS (EI5), (2) exon–intron junction reads across the 3′SS (EI3), and (3) spliced exon–exon junction reads (EE) (**Figure 2A**). The PIR of an intron was calculated by dividing the intron-retained reads by the sum of intron-retained reads and intron-skipping reads (**Figure 2A**). Constitutive introns from the annotation Wm82.a2.v1 were subjected to PIR calculations.

Alternative Splicing Analysis

Mapped reads were assembled into putative transcripts based on a reference guided assembly strategy using the single-sample transcript assembly tool StringTie v2.1.2 (Pertea et al., 2015). Multiple putative transcripts were merged into a unified set of transcripts using the meta-assembly tool TACO v0.7.3, which was considered to be superior to Cuffmerge and StringTie merge (Niknafs et al., 2017). Then, the merged transcripts were compared with the reference gene GTF file using GffCompare v0.11.2 (Pertea and Pertea, 2020). Since CB RNA was nascent RNA with no full splicing, AS analysis was based on transcripts merged from polyA RNA data. AS events were quantified based on the PSI in the program SUPPA2 (Trincado et al., 2018). Since SUPPA2 estimated the PSI based on transcript abundance, we first used salmon for alignment-free transcript abundance estimates (Patro et al., 2017). Transcripts with TPM $>$ 1 in at least three samples were used for analysis. For detection of differential splicing between two samples, we chose ΔPSI $>$ 0.1 and p-value $<$ 0.05 as cut-offs.

Detection of Differentially Expressed Genes

For detecting genes with differential expression, mapped reads in each gene were quantified using featureCounts. Then, differential gene expression was evaluated by the R package DESeq2 (Love et al., 2014). DEGs were defined by the following criteria: they had to show more than twofold up- or downregulation, and the false discovery rate (FDR)-adjusted q-value calculated by DESeq2 had to be less than 0.05. The read density for each gene was calculated by normalizing the read count to the library size and mappable length (TPM).

Gene Ontology Enrichment Analysis

Gene Ontology annotation of genes was extracted from the annotation file for Wm82.a2.v1. A hypergeometric test was explored for the statistical test, and the Benjamini and Hochberg method (1995) was used to adjust the p-value to control the FDR. All analysis was done in R software.

Detection of New Non-coding RNA Genes

To detect new ncRNA genes at the nascent RNA and polyA RNA levels, transcripts were assembled in CB RNA and polyA RNA data separately and merged by TACO as described above in the AS event analysis. Then, annotation GTF files of transcripts were compared with reference annotation GTF files using GffCompare (with the -r option). For each putative transcript, its relationship to the closest reference transcript was described by a "class code" value. For example, the code "=" indicates that the introns of a transcript completely match the introns of the reference transcript. We chose only unknown, intergenic transcripts that were assigned the code "u" and estimated their protein-coding potential by two software programs, CNCI and FEELnc (Sun et al., 2013; Wucher et al., 2017).

Reanalysis of ChIP-Seq Data of Histone Modifications

ChIP-seq raw data of histone modifications were downloaded from NCBI (**Supplementary Table 1**). The raw data were first processed with adapter removal by Cutadapt and mapping to the genome by STAR. Then, the average distribution of different histone modifications on genomic features was plotted using deepTools by normalization to histone 3 (Ramírez et al., 2016).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: NCBI website under Bioproject accession PRJNA689321.

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

AUTHOR CONTRIBUTIONS

ZD, ML, and JZ designed the research. JZ and ML performed the research. ML, HZ, FK, and BL analyzed the data. ML and ZD wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Transgenic Major Project of China (2019ZX08010003-002-015), the National Key Research and Development Program of China (2016YFD010190401), the National Natural Science Foundation of China (32090061), and Guangdong University Innovation Team Project (2019KCXTD010).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.649634/full#supplementary-material

REFERENCES

Aebi, M., Hornig, H., Padgett, R. A., Reiser, J., and Weissmann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell* 47, 555–565. doi: 10.1016/0092-8674(86)90620-3

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815. doi: 10.1038/35048692

Aslanzadeh, V., Huang, Y., Sanguinetti, G., and Beggs, J. D. (2018). Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. *Genome Res.* 28, 203–213. doi: 10.1101/gr.225615.117

Baralle, F. E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18, 437–451. doi: 10.1038/nrm.2017.27

Baurén, G., and Wieslander, L. (1994). Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* 76, 183–192. doi: 10.1016/0092-8674(94)90182-1

Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* 15, 163–175. doi: 10.1038/nrg3662

Beyer, A. L., and Osheim, Y. N. (1988). Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev.* 2, 754–765. doi: 10.1101/gad.2.6.754

Braunschweig, U., Barbosa-Morais, N. L., Pan, Q., Nachman, E. N., Alipanahi, B., Gonatopoulos-Pournatzis, T., et al. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. *Genome Res.* 24, 1774–1786. doi: 10.1101/gr.177790.114

Carrillo Oesterreich, F., Herzel, L., Straube, K., Hujer, K., Howard, J., and Neugebauer, K. M. (2016). Splicing of nascent RNA coincides with intron exit from RNA Polymerase II. *Cell* 165, 372–381. doi: 10.1016/j.cell.2016.02.045

Core, L. J., Waterfall, J. J., and Lis, J. T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848. doi: 10.1126/science.1162228

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Drexler, H. L., Choquet, K., and Churchman, L. S. (2020). Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores. *Mol. Cell* 77, 985.e8–998.e8. doi: 10.1016/j.molcel.2019.11.017

Erhard, K. F., Talbot, J. E. R. B., Deans, N. C., McClish, A. E., and Hollick, J. B. (2015). Nascent transcription affected by RNA polymerase IV in *Zea mays*. *Genetics* 199, 1107–1125. doi: 10.1534/genetics.115.174714

Gazara, R. K., de Oliveira, E. A. G., Rodrigues, B. C., Nunes da Fonseca, R., Oliveira, A. E. A., and Venancio, T. M. (2019). Transcriptional landscape of soybean (*Glycine max*) embryonic axes during germination in the presence of paclobutrazol, a gibberellin biosynthesis inhibitor. *Sci. Rep.* 9:9601. doi: 10.1038/s41598-019-45898-2

Hetzel, J., Duttke, S. H., Benner, C., and Chory, J. (2016). Nascent RNA sequencing reveals distinct features in plant transcription. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12316–12321. doi: 10.1073/pnas.1603217113

Hu, Q., Greene, C. S., and Heller, E. A. (2020). Specific histone modifications associate with alternative exon selection during mammalian development. *Nucleic Acids Res.* 48, 4709–4724. doi: 10.1093/nar/gkaa248

Jia, J., Long, Y., Zhang, H., Li, Z., Liu, Z., Zhao, Y., et al. (2020). Post-transcriptional splicing of nascent RNA contributes to widespread intron retention in plants. *Nat. Plants* 6, 780–788. doi: 10.1038/s41477-020-0688-1

Kasai, A., Kasai, K., Yumoto, S., and Senda, M. (2007). Structural features of GmIRCHS, candidate of the I gene inhibiting seed coat pigmentation in soybean: implications for inducing endogenous RNA silencing of chalcone synthase genes. *Plant Mol. Biol.* 64, 467–479. doi: 10.1007/s11103-007-9169-4

Khodor, Y. L., Rodriguez, J., Abruzzi, K. C., Tang, C. H. A., Marr, M. T., and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in *Drosophila*. *Genes Dev.* 25, 2502–2512. doi: 10.1101/gad.178962.111

Kim, S., Kim, H., Fong, N., Erickson, B., and Bentley, D. L. (2011). Pre-mRNA splicing is a determinant of histone H3K36 methylation. *Proc. Natl. Acad. Sci. U.S.A.* 108:13564. doi: 10.1073/pnas.1109475108

Kindgren, P., Ivanov, M., and Marquardt, S. (2020). Native elongation transcript sequencing reveals temperature dependent dynamics of nascent RNAPII transcription in *Arabidopsis*. *Nucleic Acids Res.* 48, 2332–2347. doi: 10.1093/nar/gkz1189

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

Laloum, T., Martín, G., and Duque, P. (2018). Alternative splicing control of abiotic stress responses. *Trends Plant Sci.* 23, 140–150. doi: 10.1016/j.tplants.2017.09.019

Li, S., Wang, Y., Zhao, Y., Zhao, X., Chen, X., and Gong, Z. (2020). Global co-transcriptional splicing in *Arabidopsis* and the correlation with splicing regulation in mature RNAs. *Mol. Plant* 13, 266–277. doi: 10.1016/j.molp.2019.11.003

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 10.1093/bioinformatics/btt656

Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R. J., Franklin, L. D., et al. (2010). An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J.* 63, 86–99. doi: 10.1111/j.1365-313X.2010.04222.x

Lin, X., Lin, W., Ku, Y.-S., Wong, F.-L., Li, M.-W., Lam, H.-M., et al. (2020). Analysis of soybean long non-coding RNAs reveals a subset of small peptide-coding transcripts. *Plant Physiol.* 182:1359. doi: 10.1104/pp.19.01324

Liu, B., Watanabe, S., Uchiyama, T., Kong, F., Kanazawa, A., Xia, Z., et al. (2010). The soybean stem growth habit gene Dt1 is an ortholog of *Arabidopsis* TERMINAL FLOWER1. *Plant Physiol.* 153, 198–210. doi: 10.1104/pp.109.150607

Liu, M., Zhu, J., and Dong, Z. (2020). Immediate transcriptional responses of *Arabidopsis* leaves to heat shock. *J. Integr. Plant Biol.* 63, 468–483. doi: 10.1111/jipb.12990

Liu, Y., Du, H., Li, P., Shen, Y., Peng, H., Liu, S., et al. (2020). Pan-genome of wild and cultivated soybeans. *Cell* 182, 162.e13–176.e13. doi: 10.1016/j.cell.2020.05.023

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

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, X., Xiong, Q., Cheng, T., Li, Q.-T., Liu, X.-L., Bi, Y.-D., et al. (2017). A PP2C-1 allele underlying a quantitative trait locus enhances soybean 100-seed weight. *Mol. Plant* 10, 670–684. doi: 10.1016/j.molp.2017.03.006

Lu, Z., Marand, A. P., Ricci, W. A., Ethridge, C. L., Zhang, X., and Schmitz, R. J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat. Plants* 5, 1250–1259. doi: 10.1038/s41477-019-0548-z

Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000. doi: 10.1126/science.1184208

Mahrez, W., Shin, J., Muñoz-Viana, R., Figueiredo, D. D., Trejo-Arellano, M. S., Exner, V., et al. (2016). BRR2a affects flowering time via FLC splicing. *PLoS Genet.* 12:e1005924. doi: 10.1371/journal.pgen.1005924

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 17, 10–12. doi: 10.14806/ej.17.1.200

McGlincy, N. J., and Smith, C. W. J. (2008). Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem. Sci.* 33, 385–393. doi: 10.1016/j.tibs.2008.06.001

Neugebauer, K. M. (2019). Nascent RNA and the coordination of splicing with transcription. *Cold Spring Harb. Perspect. Biol.* 11:a032227. doi: 10.1101/cshperspect.a032227

Niknafs, Y. S., Pandian, B., Iyer, H. K., Chinnaiyan, A. M., and Iyer, M. K. (2017). TACO produces robust multisample transcriptome assemblies from RNA-seq. *Nat. Methods* 14, 68–70. doi: 10.1038/nmeth.4078

Nojima, T., Gomes, T., Grosso, A. R. F., Kimura, H., Dye, M. J., Dhir, S., et al. (2015). Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. *Cell* 161, 526–540. doi: 10.1016/j.cell.2015.03.027

Nojima, T., Rebelo, K., Gomes, T., Grosso, A. R., Proudfoot, N. J., and Carmo-Fonseca, M. (2018). RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. *Mol. Cell* 72, 369.e4–379.e4. doi: 10.1016/j.molcel.2018.09.004

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197

Pellicer, J., and Leitch, I. J. (2020). The Plant DNA C-values database (release 7.1): an updated online repository of plant genome size data for comparative studies. *New Phytol.* 226, 301–305. doi: 10.1111/nph.16261

Pertea, G., and Pertea, M. (2020). GFF Utilities: GffRead and GffCompare. *F1000Res.* 9:ISCB Comm J-304. doi: 10.12688/f1000research.23297.1

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165. doi: 10.1093/nar/gkw257

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670

Severin, A. J., Woody, J. L., Bolon, Y.-T., Joseph, B., Diers, B. W., Farmer, A. D., et al. (2010). RNA-seq atlas of Glycine max: a guide to the soybean transcriptome. *BMC Plant Biol.* 10:160. doi: 10.1186/1471-2229-10-160

Shen, Y., Liu, J., Geng, H., Zhang, J., Liu, Y., Zhang, H., et al. (2018). *De novo* assembly of a Chinese soybean genome. *Sci. China. Life Sci.* 61, 871–884. doi: 10.1007/s11427-018-9360-0

Shen, Y., Zhou, Z., Wang, Z., Li, W., Fang, C., Wu, M., et al. (2014). Global dissection of alternative splicing in paleopolyploid soybean. *Plant Cell* 26, 996–1008. doi: 10.1105/tpc.114.122739

Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., et al. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 41:e166. doi: 10.1093/nar/gkt646

Takeshima, R., Nan, H., Harigai, K., Dong, L., Zhu, J., Lu, S., et al. (2019). Functional divergence between soybean FLOWERING LOCUS T orthologues FT2a and FT5a in post-flowering stem growth. *J. Exp. Bot.* 70, 3941–3953. doi: 10.1093/jxb/erz199

Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J. E., Nelson, R. L., et al. (2010). Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8563–8568. doi: 10.1073/pnas.1000088107

Trincado, J. L., Entizne, J. C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D. J., et al. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol.* 19:40. doi: 10.1186/s13059-018-1417-1

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., et al. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476. doi: 10.1038/nature07509

Wang, L., Cao, C., Ma, Q., Zeng, Q., Wang, H., Cheng, Z., et al. (2014). RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. *BMC Plant Biol.* 14:169. doi: 10.1186/1471-2229-14-169

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185. doi: 10.1093/bioinformatics/bts356

Wu, F., Sedivy, E. J., Price, W. B., Haider, W., and Hanzawa, Y. (2017). Evolutionary trajectories of duplicated FT homologues and their roles in soybean domestication. *Plant J.* 90, 941–953. doi: 10.1111/tpj.13521

Wucher, V., Legeai, F., Hédan, B., Rizk, G., Lagoutte, L., Leeb, T., et al. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res.* 45:e57. doi: 10.1093/nar/gkw1306

Xie, M., Chung, C. Y., Li, M. W., Wong, F. L., Wang, X., Liu, A., et al. (2019). A reference-grade wild soybean genome. *Nat. Commun.* 10:1216. doi: 10.1038/s41467-019-09142-9

Zhu, D., Mao, F., Tian, Y., Lin, X., Gu, L., Gu, H., et al. (2020). The features and regulation of co-transcriptional splicing in *Arabidopsis*. *Mol. Plant* 13, 278–294. doi: 10.1016/j.molp.2019.11.004

Zhu, J., Liu, M., Liu, X., and Dong, Z. (2018). RNA polymerase II activity revealed by GRO-seq and pNET-seq in *Arabidopsis*. *Nat. Plants* 4, 1112–1123. doi: 10.1038/s41477-018-0280-0

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.

frontiers in Plant Science

METHODS
published: 12 April 2021
doi: 10.3389/fpls.2021.634679

Rapid and Low-Input Profiling of Histone Marks in Plants Using Nucleus CUT&Tag

*Weizhi Ouyang[1], Xiwen Zhang[1], Yong Peng[1], Qing Zhang[1], Zhilin Cao[1,2], Guoliang Li[1,3] and Xingwang Li[1]**

[1] National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China, [2] Department of Resources and Environment, Henan University of Engineering, Zhengzhou, China, [3] Hubei Key Laboratory of Agricultural Bioinformatics and Hubei Engineering Technology Research Center of Agricultural Big Data, 3D Genomics Research Center, Huazhong Agricultural University, Wuhan, China

Characterizing genome-wide histone posttranscriptional modifications and transcriptional factor occupancy is crucial for deciphering their biological functions. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a powerful method for genome-wide profiling of histone modifications and transcriptional factor-binding sites. However, the current ChIP-seq experimental procedure in plants requires significant material and several days for completion. CUT&Tag is an alternative method of ChIP-seq for low-sample and single-cell epigenomic profiling using protein A-Tn5 transposase fusion proteins (PAT). In this study, we developed a nucleus CUT&Tag (nCUT&Tag) protocol based on the live-cell CUT&Tag technology. Our results indicate that nCUT&Tag could be used for histone modifications profiling in both monocot rice and dicot rapeseed using crosslinked or fresh tissues. In addition, both active and repressive histone marks such as H3K4me3 and H3K9me2 can be identified using our nCUT&Tag. More importantly, all the steps in nCUT&Tag can be finished in only 1 day, and the assay can be performed with as little as 0.01 g of plant tissue as starting materials. Therefore, our results demonstrate that nCUT&Tag is an efficient alternative strategy for plant epigenomic studies.

Keywords: CUT&Tag, chromatin profiling, histone modification, ChIP-seq, native nucleus, nCUT&Tag

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Roger Deal,
Emory University, United States
Frederic Berger,
Gregor Mendel Institute of Molecular Plant Biology (GMI), Austria

***Correspondence:**
Xingwang Li
xingwangli@mail.hzau.edu.cn

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

Received: *28 November 2020*
Accepted: *19 March 2021*
Published: *12 April 2021*

Citation:
Ouyang W, Zhang X, Peng Y, Zhang Q, Cao Z, Li G and Li X (2021) Rapid and Low-Input Profiling of Histone Marks in Plants Using Nucleus CUT&Tag. Front. Plant Sci. 12:634679. doi: 10.3389/fpls.2021.634679

INTRODUCTION

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is an efficient method for profiling histone modifications and transcription factor-binding sites (Johnson et al., 2007). In the standard ChIP-seq assay for plants (Kaufmann et al., 2010), formaldehyde-fixed nuclei are isolated and sonicated. Thereafter, the fragmented chromatin is prepared for immunoprecipitation and the ChIP DNA is purified and fragmented for sequencing library preparation. The standard plant ChIP-seq assays are complex, requiring large numbers of input cells/tissues and lasting several days from sample fixation to the sequencing-ready library. To improve chromatin profiling efficiency and save experiment time, Zhao et al. (2020) developed an enhanced ChIP-seq (eChIP-seq) protocol

with modifications to the standard ChIP-seq. In eChIP-seq, the homogenate chromatin lysates are directly sonicated without nuclei purification steps. Hence, eChIP-seq considerably boosts chromatin extraction efficiency and saves a significant amount of time compared to the traditional ChIP-seq method (Zhao et al., 2020).

Recently, CUT&RUN and CUT&Tag have been developed by fusing protein A (PAT) with micrococcal nuclease and Tn5 transposase, respectively, to study chromatin state profiling using low-input samples or single live cells (Skene et al., 2018; Kaya-Okur et al., 2019). With CUT&Tag (Kaya-Okur et al., 2019), Tn5 transposase, in fusion to PAT, is tethered at specific genomic regions through the affinity of PAT to interested antibodies. Then, activation of Tn5 generates chromatin fragments for direct PCR amplification. Compared to ChIP-seq, CUT&Tag omits many steps, such as sonication, chromatin immunoprecipitation, and complicated library preparation (including DNA end repair, A-tailing, adapter ligation, and PCR enrichment). Hence, CUT&Tag enables the processing of chromatin profiling with low-input samples or even single cells and manipulation of the entire experimental procedure in only 1 day. Moreover, the PAT-based chromatin profiling strategies eliminate the requirement of the sonication and immunoprecipitation steps, enabling high-throughput identification of histone modifications at single-cell levels (Carter et al., 2019; Kaya-Okur et al., 2019; Wang et al., 2019). Most recently, Tao et al. (2020) profiled the H3K4me3 modification in cotton with high resolution and low background noise using CUT&Tag. However, the cotton CUT&Tag assays still required a significant quantity of input tissue and were time-consuming (2–3 days).

In this study, we employed our previously reported protocols for rapid and efficient nuclei isolation and developed a nucleus CUT&Tag (nCUT&Tag) protocol with protein G-Tn5 (PGT) for rapid and low-input histone modification profiling using crosslinked and fresh plant tissue. Our results showed that nCUT&Tag is an alternative strategy of ChIP-seq for fast and low-input profiling both active and repressive histone marks with crosslinked or fresh tissues from the monocots or dicots.

MATERIALS AND EQUIPMENT

Plant Materials, Growth Conditions, and Sample Collection

The rice cultivar from the Xian group (known as *Oryza sativa* L. ssp. *indica*), Minghui 63 (MH63), was grown in a growth chamber with the day/night cycle set at 14/10 h and a temperature of 32/28°C. The 15-day-old seedlings were collected for fresh nCUT&Tag, or crosslinked with 1% formaldehyde solution for crosslinking nCUT&Tag. A rice hybrid MHNip (MH63 × Nipponbare) was used for panicles collection. MHNip was planted in the field of Huazhong Agricultural University, Wuhan, China, and grew under normal agricultural conditions. Young panicles with 2.5-4 cm in length were collected and dual-crosslinked with 1% formaldehyde and EGS. The *Brassica napus* cultivar 2063A was grown in the growth chamber. Young leaves of 21-day-old 2063A seedlings were harvested and crosslinked with 1% formaldehyde solution.

Regents and Equipment

1. Antibodies against proteins of interest:
   Anti-H3K4me3 (Abclonal, A2357; 1 mg/ml)
   Anti-H3K9me2 (Abcam, ab1220; 1 mg/ml)
2. Protein G-Tn5 fusion protein (Vazyme, cat. no. S602)
3. Phosphate-buffered saline (PBS) (Ambion, cat. no. AM9625)
4. Formaldehyde (37%; EMD Millipore, cat. no. 344198-250ML)
5. Ethylene glycol bis (succinimidyl succinate) (EGS; Thermo Fisher Scientific, cat. no. 21565)
6. Glycine (Sigma–Aldrich, cat. no. G8898-500G)
7. Sodium deoxycholate (Sigma–Aldrich, cat. no. 30970-100G)
8. Triton X-100, molecular biology grade (Promega, cat. no. H5141)
9. Tween 20 for molecular biology, viscous liquid (Sigma–Aldrich, cat. no. P9416-100ML)
10. HEPES buffer (1 M, pH 7.3, Fisher Scientific, cat. no. BP299-1)
11. NaCl solution (500 ml, 5.0M, Ambion, cat. no. AM9759)
12. Spermidine (Sigma, cat. no. S2501-1G) 2 M
13. Complete Protease Inhibitor (Roche, cat. no. 5056489001)
14. Nuclease-Free Water (1000 ml; Ambion cat. no. 4387936)
15. EDTA (pH 8.0, 0.5 M, 500 ml; Ambion, cat. no. AM9261)
16. Bovine serum albumin (BSA) (Sigma, cat. no. A1933-100G)
17. $MgCl_2$ (1 M, 100 ml; Ambion, cat. no. AM9530G)
18. Sodium dodecyl sulfate (SDS, wt/vol 10%; Ambion, cat. no. AM9822)
19. Proteinase K solution (Life Technologies, cat. no. AM2548)
20. Phenol:chloroform:IAA 25:24:1 (Ambion, cat. no. AM9730)
21. GlycoBlue (Life Technologies, cat. no. AM9516)
22. Isopropanol (Sigma–Aldrich, cat. no. I-9516-500ml)
23. Sodium acetate (Ambion, cat. no. AM9740)
24. Absolute ethanol (500 ml; Sigma–Aldrich, cat. no. E7023)
25. MinElute PCR purification kit (Qiagen, cat. no. 28004)
26. TruePrep DNA Library Prep Kit V2 for Illumina (Vazyme cat. no. TD501)
27. AMPure XP beads (60 ml; Beckman, cat. no. A63881)
28. Buffer EB (250 ml; Qiagen, cat. no. 19086)
29. Dynabeads Protein G for immunoprecipitation (50 ml; Life Technologies, cat. no. 10009D)
30. Qubit 3.0 Fluorometer (Invitrogen, cat. no. Q33216)
31. Bio-Rad C1000 Thermal Cycler (Bio-Rad, cat. no. 185-1148EDU)
32. Centrifuge (Eppendorf 5810R, Swing-bucket Rotor with 15- and 50-ml Buckets, cat. no. 22628180)
33. Bioruptor Plus (UCD-300; Diagenode, cat. no. B01020001).

Regent Setup

1. Wash Buffer (50 ml): Add 1 ml HEPES buffer (1 M, pH 7.5), 1.5 ml NaCl (5 M), and 12.5 μl spermidine (2 M) together and fill with distilled water to a final volume of 50 ml. Dissolve one tablet of Complete Protease Inhibitor in the buffer before use. Store the buffer at 4°C for up to 1 week.

2. Antibody Buffer (250 μl): Mix 1 μl EDTA (pH 8.0, 0.5 M) and 0.8 μl BSA (30%) with 250 μl Wash Buffer and chill on ice until use.
3. Transposase Incubation Buffer (50 ml): Add 1 ml HEPES buffer (1 M, pH 7.5), 3 ml NaCl (5 M), and 12.5 μl spermidine (2 M) together and bring the final volume to 50 ml with distilled water. Store the buffer at 4°C for up to 1 week. Dissolve one tablet of Complete Protease Inhibitor in the buffer before use.
4. Tagmentation Buffer (300 μl): Mix 300 μl Transposase Incubation Buffer and 3 μl $MgCl_2$ (1 M) together.
5. Buffer S (500 ml): Add 25 ml HEPES buffer (1 M, pH 7.5), 15 ml NaCl (5 M), 1 ml EDTA (0.5 M), 5 ml Triton X-100, 5 ml sodium deoxycholate (10%), and 50 ml SDS (10%) together; mix the solution well and bring the final volume to 500 ml with distilled water. Sterile filtrate and store at room temperature for up to 6 months.
6. Buffer F (500 ml): Add 25 ml HEPES buffer (1 M, pH 7.5), 15 ml NaCl (5 M), 1 ml EDTA (0.5 M), 5 ml Triton X-100, and 5 ml sodium deoxycholate (10%) together; mix the solution well and bring the final volume to 500 ml with distilled water. Sterile filtrate and store at 4°C for up to 6 months.
7. Binding Buffer (10 ml): Add 200 μl HEPES buffer (1 M, pH 7.5), 100 μl KCl (1 M), 10 μl $CaCl_2$ (1 M), 10 μl $MnCl_2$ (1 M) together and bring the final volume to 10 ml with distilled water. Store at 4°C for up to 6 months.

METHODS

Nuclei Isolation

Formaldehyde-fixed nuclei are isolated according to our previously reported protocols (**Figure 1A**) (Zhao et al., 2020). Briefly, 0.1 or 0.01 g of crosslinked tissue is ground to fine powders in liquid nitrogen. The powder is suspended with 300 μl Buffer S and lyzed at 4°C for 30 min with rotation. Then the 300 μl lysates are mixed with 1.2 ml Buffer S and lyzed at 4°C for 15 min with rotation. Finally, the homogenate lysates are centrifuged at 1000 *g* for 10 min at 4°C, and the nuclei are collected.

The native nuclei from fresh tissue, as well as formaldehyde-fixed nuclei from crosslinked tissue, can be isolated following a simple and fast strategy (**Figure 1B**) (Sun et al., 2020). The plant tissue is chopped thoroughly to complete homogeneity in a plastic petri dish with 1 ml 1 × PBS (containing protease inhibitor) on ice. The homogenate is filtered twice through a layer of Miracloth. The nuclei are isolated by centrifuging the filtrate in a swinging bucket rotor at 1000 *g* for 10 min at 4°C.

The collected nuclei are stained with DAPI and observed under a fluorescence microscope. All eChIP-seq libraries are prepared following our reported protocols with Buffer S/F isolated nuclei (Zhao et al., 2020). nCUT&Tag starts with fixed or native nuclei, followed by subsequent antibody binding to proteins of interest, PGT binding to antibodies, tagmentation, DNA purification, library preparation, and sequencing (**Figure 2**). The following procedures are a detailed introduction of the nCUT&Tag protocol.

Procedures for nCUT&Tag

Antibody Binding to Target Protein

1. Wash the nuclei pellet twice with 500 μl ice-cold Wash Buffer. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C; discard Wash Buffer.
2. Resuspend the nuclei pellet in 200 μl ice-cold Antibody Buffer. Divide into two 1.5 ml tubes with 100 μl each.
3. Add 1–5 μg antibody and IgG to the two 100 μl suspensions, respectively.
4. Incubate at 4°C for 2 h with rotation.

PGT Binding to Antibody

1. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C. Discard the Antibody Buffer.
2. Wash the nuclei pellet with 800 μl ice-cold Wash Buffer. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C; discard Wash Buffer.
3. Repeat Step 2 twice.
4. Mix 100 μl Transposase Incubation Buffer and 0.58 μl assembled PGT (final concentrate: 0.04 μM). Resuspend the nuclei pellet in the 100 μl transposase mixture with gentle vortexing.
5. Incubate at 4°C for 1 h with rotation.

Tagmentation

1. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C. Discard the supernatant.
2. Wash the nuclei pellet with 800 μl ice-cold Transposase Incubation Buffer. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C; discard Transposase Incubation Buffer.
3. Repeat Step 2 twice.
4. Mix 300 μl Transposase Incubation Buffer and 3 μl $MgCl_2$ together and resuspend the nuclei pellet.
5. Incubate at 37°C for 1 h.

DNA Purification

1. Add 10 μl EDTA (0.5 M) and 3 μl SDS (10% wt/vol) to stop tagmentation.
   Note: for fresh tissue, the Qiagen MinElute PCR purification kit (Qiagen, cat. no. 28004) is optional for DNA purification without prior reverse crosslinking. It saves much time.
2. Add 2.5 μl proteinase K solution and incubate at 50°C for 1 h to release DNA.
3. Add an equal volume of phenol–chloroform–isoamyl alcohol (pH 7.9) to the tagmentation product and mix vigorously.
4. Spin MaXtract High Density tubes at 16,000 *g* for 2 min at room temperature. Transfer the mixture in Step 3 to the centrifuged MaXtract High Density tubes and centrifuge at 16,000 *g*, at room temperature for 5 min.
5. Transfer upper aqueous phase above the gel matrix to fresh 1.5-ml tubes; add 30 μl 3 M sodium acetate (pH 5.5), 2 μl GlycoBlue, and 330 μl isopropanol and mix them well.

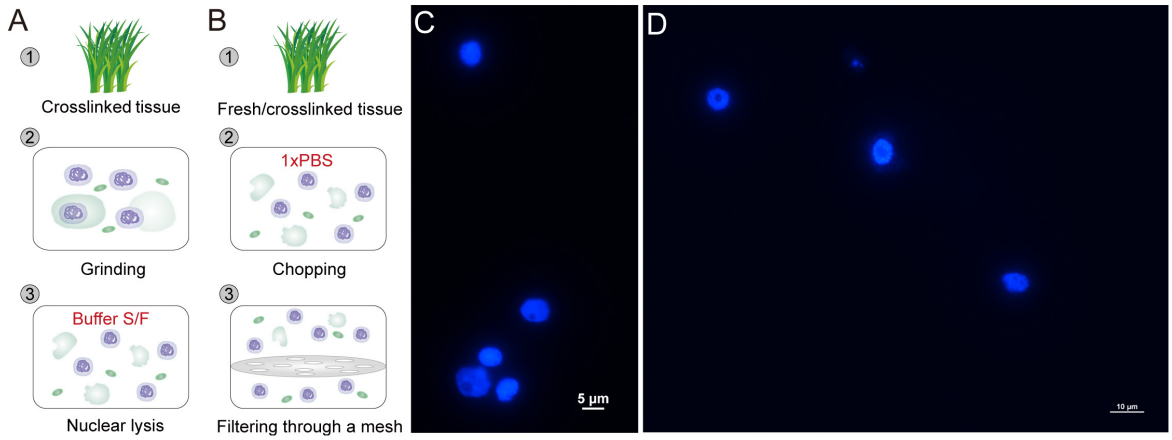

FIGURE 1 | Rapid and efficient isolation of high-integrity nuclei. **(A,B)** Two strategies for nuclei isolation with crosslinked or fresh tissue. Crosslinked tissue is ground to fine powder and lyzed with Buffer S and Buffer F **(A)**. Crosslinked or fresh tissue is chopped to homogenate lysates in PBS and filtered through a mesh **(B)**. The released nuclei are collected by centrifugation. The nuclei isolated with Buffer S/F **(C)** or PBS **(D)** are stained with DAPI and observed under a fluorescence microscope.

6. Incubate and cool down at −80°C for 30 min.
7. Centrifuge at 16,000 *g* for 20 min at 4°C.
8. Wash the pellet twice with 1 ml 75% ethanol.
9. Air-dry the DNA pellet and dissolve the DNA with 50 μl QIAGEN Buffer EB.
10. Quantitate DNA using Qubit3.0 according to the manufacturer's instructions.

PCR Enrichment, Library DNA Purification, and Sequencing

50–100 ng PGT cut DNA is used for direct PCR enrichment according to the TruePrep DNA Library Prep Kit manual (Vazyme, cat. no. TD501). The PCR is performed for 13–15 cycles. PCR enriched library DNA is purified and size-selected with AMPure XP beads, and sequenced with pair-end 150 at the Illumina HiSeq2500 or HiSeq X Ten sequencing platforms.

Procedures for Low-Input nCUT&Tag

Collect 0.1 or 0.01 g of crosslinked tissue and grind to fine powders in liquid nitrogen. Resuspend the powder with 300 μl Buffer S and lyze at 4°C for 30 min with rotation. Mix the 300 μl lysates with 1.2 ml Buffer S and lyze at 4°C for another 15 min with rotation. Centrifuge the homogenate lysates at 1000 *g* for 10 min at 4°C and collect the nuclei.

Binding Nuclei to Concanavalin A-Coated Magnetic Beads (Con-A Beads)

1. Wash the nuclei pellet twice with 500 μl ice-cold Wash Buffer. Centrifuge in a swinging bucket rotor at 600 *g* for 3 min at 4°C; discard Wash Buffer.
2. Wash 20 μl Con-A beads with 500 μl Binding Buffer twice to activate Con-A beads. Place the tube on a magnet stand and remove the liquid.
3. Resuspend Con-A beads with 100 μl Binding Buffer. Add the activated beads to isolated nuclei and incubate the mixture at 4°C for 15 min.

Antibody Binding to Target Protein

1. Discard the liquid and collect nuclei by a magnetic stand.
2. Resuspend the nuclei in 200 μl ice-cold Antibody Buffer. Divide into two 1.5 ml tubes with 100 μ l each.
3. Add 1 μg antibody and IgG to the two 100 μl suspensions, respectively.
4. Incubate at 4°C for 2 h with rotation.

PGT Binding to Antibody

1. Discard the liquid and collect nuclei by a magnetic stand.
2. Wash the nuclei with 800 μl ice-cold Wash Buffer. Discard the liquid and collect nuclei by a magnetic stand.
3. Repeat Step 2 twice.
4. Mix 100 μl Transposase Incubation Buffer and 0.58 μl assembled PGT (final concentrate: 0.04 μM). Resuspend the nuclei in the 100 μl transposase mixture with gentle vortexing.
5. Incubate at 4°C for 1 h with rotation.

Tagmentation

1. Discard the liquid and collect nuclei by a magnetic stand.
2. Wash the nuclei with 800 μl ice-cold Transposase Incubation Buffer. Discard the liquid and collect nuclei by a magnetic stand.
3. Repeat Step 2 twice.
4. Mix 300 μl Transposase Incubation Buffer and 3 μl MgCl2 together and resuspend the nuclei.
5. Incubate at 37°C for 1 h.

DNA Purification and Library Preparation

Purify the tagmented DNA and prepare the sequencing library following the procedures as described in the nCUT&Tag protocol above.

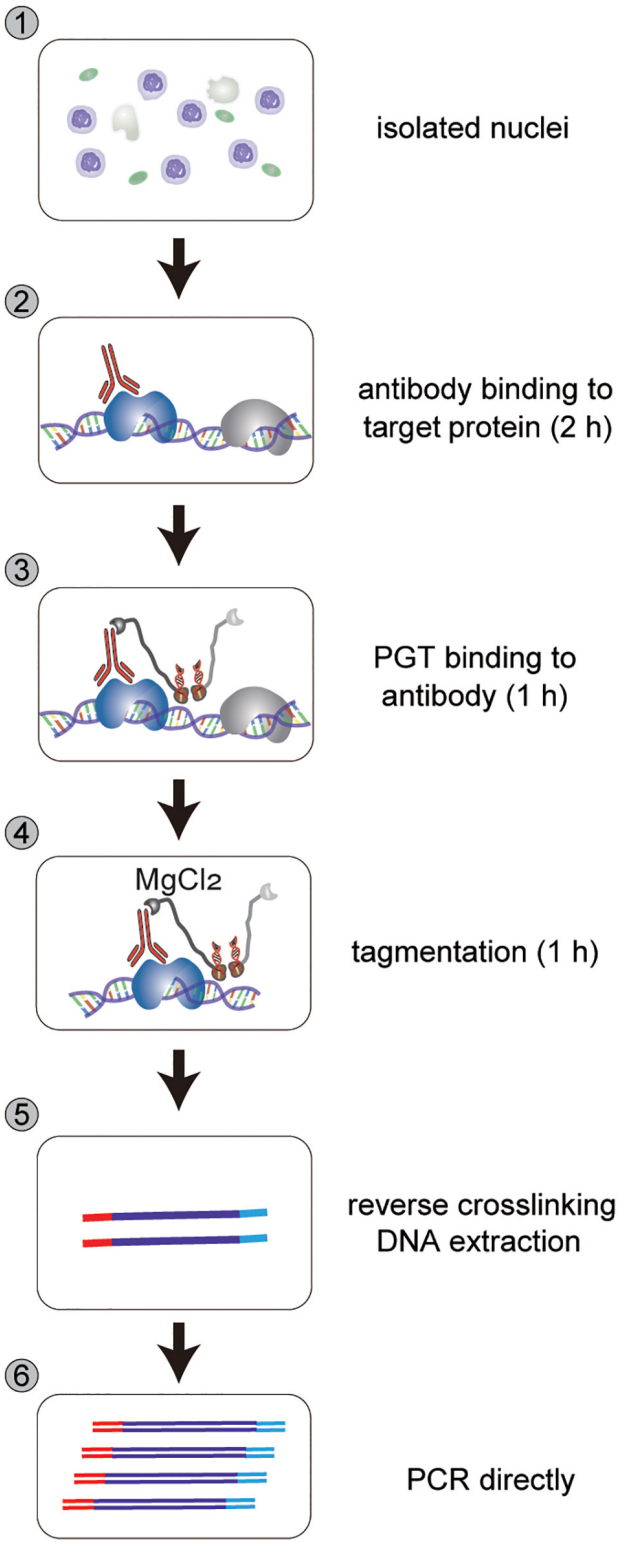

FIGURE 2 | Workflow of nCUT&Tag. nCUT&Tag starts with isolated nuclei, followed by antibody binding to target protein for 2 h, PGT binding to antibody for 1 h, tagmentation for 1 h, reverse crosslinking, and direct PCR for library DNA enrichment. For fresh tissue, reverse crosslinking (step 5) can be omitted. Tagmentation DNA is purified directly using a DNA Purification kit.

Bioinformatic Analysis

Trimmomatic (v0.32) (Bolger et al., 2014) is used to remove low-quality reads and to trim low-quality bases as well as adapters, with the following parameters: "ILLUMINACLIP:/adapters/TruSeq3-PE.fa:2:30:10:8:True SLIDINGWINDOW:4:15 MINLEN:50 HEADCROP:10 LEADING:5 TRAILING:5." Trimmed reads are aligned to the MH63 reference genome (MHRS2) (Zhang et al., 2016) or *B. napus* reference genome (Chalhoub et al., 2014) using BWA (v0.7.17) mem with default settings (Li and Durbin, 2009). Then alignments with MAPQ < 30 and duplicated reads are discarded using samtools (v1.9) (Li et al., 2009). Peak calling for H3K4me3 uses macs2 (v2.1.1) with the following parameters: macs2 callpeak -t treat_bam -c control_bam -f BAMPE -B -q 0.05 -g 3.6e + 8 (-g 1.1e + 9 for *B. napus*) (Zhang et al., 2008). Broad peak calling for H3K9me2 is similar to the narrow H3K4me3 peak calling with an additional parameter –broad. Scatterplots, correlation plots, and the signal heatmaps are created using deepTools (v2.5.3) (Ramirez et al., 2014) as previously described (Zhao et al., 2020). Annotation of peaks is performed using homer (v4.11) annotatePeaks.pl with default parameters (Heinz et al., 2010). To compare the robust profiles of nCUT&Tag and eChIP-seq, we randomly extracted 500-K, 1-M, 2-M, 4-M, 8-M, 16-M, and 24-M valid clean reads from each samples to call peaks and calculate fraction of reads in peaks (FRiP) values as described in Kaya-Okur et al. (2019).

RESULTS

Rapid and Efficient Isolation of High-Quality Nuclei

CUT&RUN and CUT&Tag were initially developed with human live cells (Skene et al., 2018; Kaya-Okur et al., 2019). With digitonin treatment, the membrane of cell and nucleus was permeabilized so that antibody and PAT/PGT can spread into the nuclei without compromising nuclear integrity. Since cell walls are present in plant cells, it is difficult for antibody and PGT to penetrate the cells and nuclei. As an alternative, the previously reported CUT&RUN and CUT&Tag in plants started with isolated nuclei rather than live cells (Zheng and Gehring, 2019; Tao et al., 2020). However, the nuclei isolation protocols require significant material and much time for completion because of the multiple purification steps. Here, we employed two simple protocols for fast nuclei isolation with formaldehyde-fixed tissue or fresh tissue (**Figures 1A,B**). For the Buffer S/F method (**Figure 1A**), formaldehyde-fixed tissue is ground into fine powder in liquid nitrogen and lyzed with Buffer S and F. The released nuclei are collected by centrifuged at 1000 *g* for 10 min at 4°C. In the PBS strategy (**Figure 1B**), fixed or fresh tissue is chopped to complete homogeneity in a plastic petri dish with 1 × PBS (containing protease inhibitor) on ice and filtered twice through a layer of Miracloth. The nuclei are isolated by centrifuging the filtrate in a swinging bucket rotor at 1000 *g* for 10 min at 4°C. All the procedures can be finished within tens of minutes. Both the two strategies isolate high-integrity nuclei

(**Figures 1C,D**). It is worth noting that the PBS strategy is a mild and fast method for nuclei isolation. It can be used for isolating nuclei from both cryopreserved crosslinked tissues and fresh tissues, while the Buffer S/F, which contains high-concentrate SDS, is a relatively harsh strategy that may be not suitable for fresh tissues. However, compared to the PBS strategy, which may lost too much nuclei (more than 80%) during the mesh-filtering step, the Buffer S/F method is a better choice for isolating high-yield and high-quality nuclei from low-input crosslinked tissues (Zhao et al., 2020).

nCUT&Tag for Rapid Chromatin Profiling With Crosslinked Tissue

The isolated nuclei were then directly incubated with antibodies and subsequently with PGT fusion protein (**Figure 2**). The PGT tagmentation reaction was activated by adding divalent magnesium ions to the incubated nuclei, and DNA fragmentation reactions occurred around the histone modification sites. Finally, the fragmented DNA was purified for sequencing library preparation.

Using nCUT&Tag, we first profiled the active chromatin features with H3K4me3 antibody using formaldehyde-fixed rice young panicles (**Figure 3A**). We performed two biological replicates of nCUT&Tag with ∼1 g of finely ground panicle powder. The nuclei were released by adding buffer S and buffer F (Zhao et al., 2020). The homogenate lysates were then centrifuged for 3 min; the nuclei pellets were used to conduct nCUT&Tag. The two replicates showed a high degree of reproducibility (r = 0.98, Spearman's correlation) and a high correlation with the H3K4me3 eChIP-seq data (r = 0.92, Spearman's correlation) (**Figure 3B** and **Supplementary Figure 1A**). The two nCUT&Tag replicates totally identified 31,483 high-confidence H3K4me3 peaks in rice young panicles (31,436 and 27,857, respectively) (**Supplementary Table 1**); among the 31,483 peaks, approximately 80% (25,497 peaks) were also detected by the eChIP-seq experiments (**Figures 3C,D**). Significantly, 5986 peaks were detected by nCUT&Tag only, while 4460 peaks were detected by eChIP-seq only (**Figures 3A,C**), indicating that the two different strategies might have distinct advantages in detecting specific histone modification sites.

The H3K4me3 peaks mainly enriched around the transcription start sites (TSS) (**Supplementary Figure 3A**), consistent with our previous eChIP-seq data (Zhao et al., 2020). In addition, peak annotation showed that more than 50% H3K4me3 peaks were distributed at gene promoters, the first exons, and the first introns; about 20% were distributed at transcription end sites (TES) and intergenic regions; the remainders were distributed across other exons and other introns (**Supplementary Table 2**). The distribution profiles of nCUT&Tag peaks showed high consistency with that of eChIP-seq peaks (**Supplementary Figure 4**).

Enhanced ChIP-seq is an efficient protocol in profiling histone marks. It was used to map rice and *B. napus* reference epigenomes with considerably low background noise (Zhang et al., 2020; Zhao et al., 2020). To compare the signal-to-noise ratio of nCUT&Tag relative to eChIP-seq, we downsampled the sequencing reads at varying depth from 1-M to 24-M. Then we called peaks and calculated FRiP values under the same sequencing depth (**Supplementary Table 3**). The results showed that eChIP-seq data exhibited higher signal-to-noise ratio than nCUT&Tag. However, using the 8-M nCUT&Tag reads, we called 27,043 peaks, which were nearly as much as that from 16-M eChIP-seq reads (27,659 peaks) (**Supplementary Table 3**). Our results indicated that nCUT&Tag showed a little bit higher background noise than the eChIP-seq protocol, but nCUT&Tag detected comparable peaks with much less sequencing reads.

nCUT&Tag for Profiling Both Active and Repressive Histone Marks With Fresh Tissue

Furthermore, we conducted H3K4me3 nCUT&Tag for native nuclear chromatin profiling with fresh rice seedlings. We isolated native nuclei, according to Sun et al. (2020). A few pieces of young leaves were chopped into homogenate lysates in PBS buffer. The lysates were filtered twice through a mesh; nuclei were collected by centrifugation and used to perform nCUT&Tag. After stopping the tagmentation reaction, the fragmented DNA was directly purified following the procedure reported for ATAC-seq (Buenrostro et al., 2013; Sun et al., 2020) using a Qiagen MinElute kit (QIAGEN, cat. no. 28004) that eliminates the reverse-crosslinking steps and is a rapid DNA purification protocol.

The fresh nCUT&Tag showed a high correlation with the fixed H3K4me3 eChIP-seq data (r = 0.92, Spearman's correlation) (**Supplementary Figures 1B**, **2A**). The two replicates totally called 26,543 peaks (21,203 and 23,545, respectively, **Supplementary Figure 4A** and **Supplementary Table 1**). Among the 26,543 peaks, 24,913 (93.86%) were also detected by eChIP-seq. Strikingly, 1730 peaks were detected by fresh nCUT&Tag only, while 5485 peaks were detected by fixed eChIP-seq only (**Figures 4B,D**). In fact, there were slight signal enrichment in nCUT&Tag libraries at the 5485 eChIP-seq unique peak regions (**Figure 4D**). A possible explanation for that many peaks were only detected by eChIP-seq may be the lower sequencing depth of the nCUT&Tag libraries relative to the eChIP-seq data (**Supplementary Table 1**). The fresh nCUT&Tag signal showed similar enrichment as that of crosslinking eChIP-seq, mainly around the TSS (**Supplementary Figure 3B**). The fresh nCUT&Tag peak distribution profiles were also similar to that of crosslinking eChIP-seq (**Supplementary Figure 4** and **Supplementary Table 2**). These results suggest that the nCUT&Tag method could be applied for mapping active histone modifications with native nuclei.

Meanwhile, we performed H3K4me3-associated nCUT&Tag with crosslinked seedlings to compare with the fresh nCUT&Tag data (**Supplementary Figure 5A**). They showed a high correlation between the fixed and fresh nCUT&Tag (r = 0.89, Spearman's correlation) (**Supplementary Figure 5B**). We detected 21,445 H3K4me3 peaks in crosslinked seedlings (**Supplementary Table 1**), among which 77% (16,468 peaks) were also detected in fresh seedlings by nCUT&Tag (**Supplementary Figures 5C,D**). Strikingly, about 10,175 peaks (∼38%) were

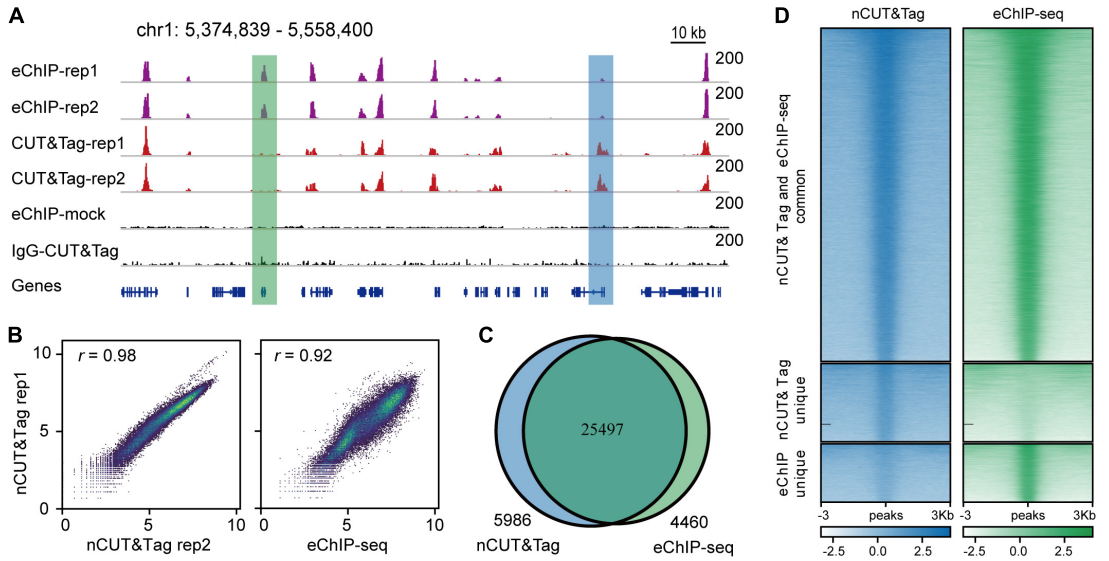

FIGURE 3 | nCUT&Tag for fast chromatin profiling with crosslinked tissue. **(A)** Representative H3K4me3 landscapes across chr1:5,374,839–5,558,400 of the rice genome generated by eChIP-seq and nCUT&Tag. The green and blue boxes show peaks detected only by eChIP-seq and nCUT&Tag, respectively. **(B)** Scatter plots showing the Spearman's correlations for the two H3K4me3 nCUT&Tag replicates (left), and between the nCUT&Tag and eChIP-seq data (right). **(C)** Venn diagram showing the overlap of H3K4me3 peaks detected by eChIP-seq and nCUT&Tag. **(D)** Comparison of the H3K4me3 eChIP-seq and nCUT&Tag signals.

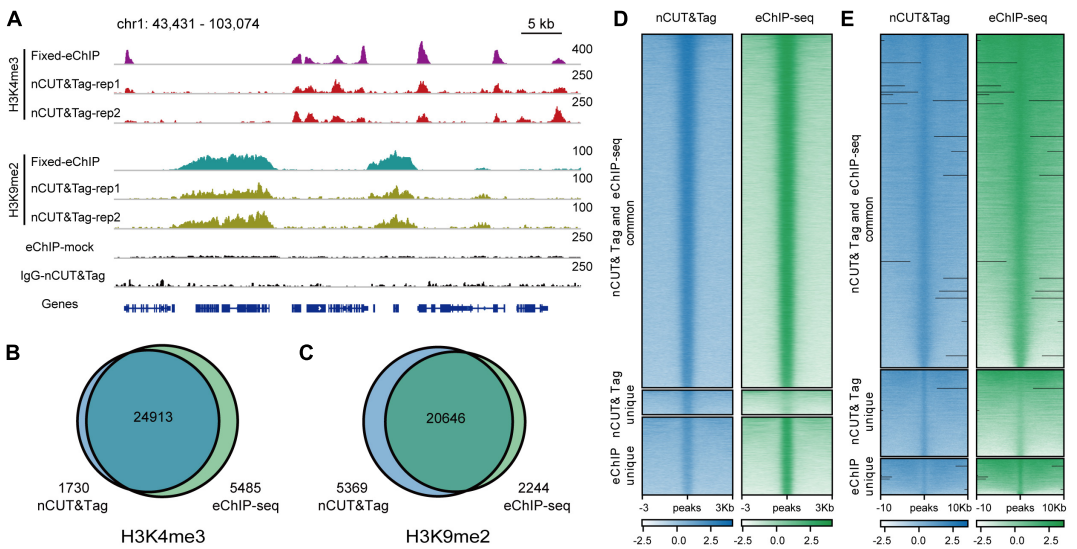

FIGURE 4 | nCUT&Tag for chromatin landscape profiling with non-crosslinked tissue. **(A)** Genome browser screenshot showing H3K4me3 and H3K9me2 nCUT&Tag data for fresh rice seedlings. The H3K4me3 and H3K9me2 eChIP-seq data were generated with crosslinked seedlings. **(B)** Venn diagram showing the overlap of H3K4me3 peaks detected by fresh nCUT&Tag and crosslinked eChIP-seq. **(C)** Venn diagram showing the overlap of H3K9me2 peaks detected by fresh nCUT&Tag and crosslinked eChIP-seq. Comparison of nCUT&Tag and eChIP-seq signals for H3K4me3 **(D)** and H3K9me2 **(E)**.

exclusively detected in fresh tissues, suggesting that crosslinking might underpresent the detection of histone modifications. For cryopreserved seedlings, the crosslinked nuclei might need to be pre-opened with Hypotonic Buffer containing SDS, as described in CoBATCH and itChIP-seq (Ai et al., 2019; Wang et al., 2019), to capture much more signals.

H3K9me2, which shows a broad-peak profile in the rice genome, is a repressive histone mark associated with closely compacted heterochromatin (Zhao et al., 2020). To test whether nCUT&Tag can be used to characterize repressive chromatin features, we conducted another nCUT&Tag procedure with fresh rice seedlings to profile H3K9me2 histone modification.

The H3K9me2 nCUT&Tag showed a high correlation with our eChIP-seq data (r = 0.95, Spearman's correlation) (**Supplementary Figures 1C**, **2B**). The two biological replicates called 24,382 and 22,142 H3K9me2 peaks, respectively

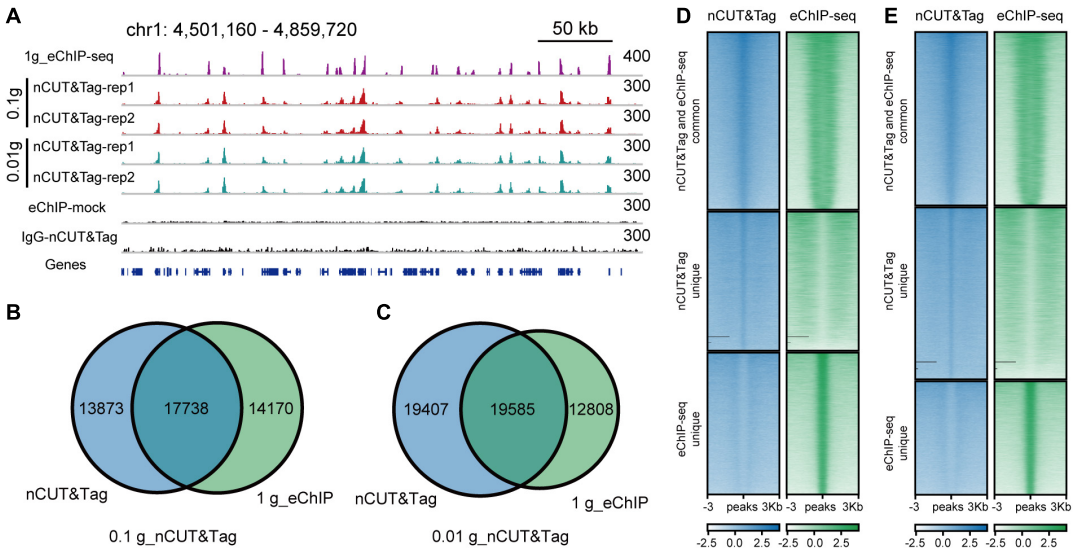

FIGURE 5 | nCUT&Tag for chromatin landscape profiling with low-input samples. **(A)** Genome browser screenshot showing H3K4me3 nCUT&Tag data for low-input rice seedlings. The H3K4me3 eChIP-seq data were generated with 1-g crosslinked seedlings. **(B)** Venn diagram showing the overlap of H3K4me3 peaks detected by 0.1-g nCUT&Tag and crosslinked eChIP-seq. **(C)** Venn diagram showing the overlap of H3K4me3 peaks detected by 0.01-g nCUT&Tag and crosslinked eChIP-seq. **(D)** Comparison of H3K4me3 signals between 0.1-g nCUT&Tag and 1-g eChIP-seq. **(E)** Comparison of H3K4me3 signals between 0.01-g nCUT&Tag and 1-g eChIP-seq.

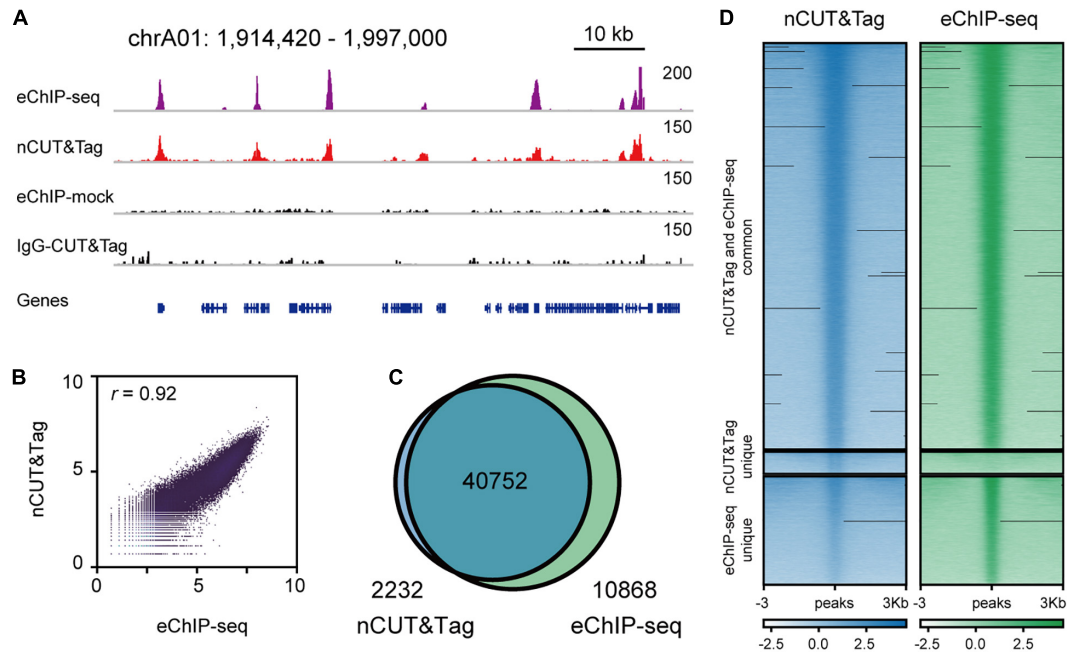

FIGURE 6 | nCUT&Tag for fast chromatin profiling in *Brassica napus*. **(A)** Representative H3K4me3 landscapes across chrA01:1,914,420–1,996,975 of the *Brassica napus* genome generated by nCUT&Tag and eChIP-seq. **(B)** Scatter plots showing the Spearman's correlation between the nCUT&Tag and eChIP-seq data. **(C)** Venn diagram showing the overlap of H3K4me3 peaks detected by nCUT&Tag and eChIP-seq. **(D)** Comparison of the H3K4me3 nCUT&Tag and eChIP-seq signals.

(**Figure 4A** and **Supplementary Table 1**). Among the 26,015 peaks identified by the two nCUT&Tag replicates, about 80% (20,646 peaks) were also detected by the eChIP-seq; 5369 peaks were detected by the fresh nCUT&Tag only, while 2244 were detected by eChIP-seq only (**Figures 4C,E**). The peak distribution showed a considerably consistency between the H3K9me2 nCUT&Tag and eChIP-seq results, with approximately 40% distributing at intergenic regions (**Supplementary Figure 4** and

Supplementary Table 2). We also compared the signal levels under the same sequencing depth between the nCUT&Tag and eChIP-seq libraries. The 8-M nCUT&Tag reads called 17,375 H3K9me2 peaks, which were almost as much as the 16-M reads from the eChIP-seq libraries (17,810 peaks) and even a little bit less than that from the 24-M eChIP-seq reads (18,954 peaks) (**Supplementary Table 3**). The results indicate that 8-M clean reads from nCUT&Tag provide comparable signals to the 16-M and even 24-M eChIP-seq reads.

Taken together, nCUT&Tag is a versatile method that can be used for global profiling of both active and repressive histone modifications in rice.

nCUT&Tag for Efficient Chromatin Profiling With Low-Input Samples

The standard ChIP-seq assay requires significant material (∼1 g). To test whether nCUT&Tag could profile histone modifications using low-input samples, we performed H3K4me3 nCUT&Tag with 0.1 and 0.01 g of crosslinked seedlings (**Figure 5A**). To avoid too much loss of nuclei in the centrifuge steps, here we use concanavalin A-coated magnetic beads for buffer exchange as an alternative strategy. The low-sample nCUT&Tag showed high correlations with both regular nCUT&Tag and eChIP-seq (the Spearman's correlations varied from 0.89 to 0.93), and a high degree of consistency between the 0.1-g nCUT&Tag and 0.01-g nCUT&Tag (r = 1.00, Spearman's correlation) (**Supplementary Figures 1B**, **2A**). The 0.1- and 0.01-g nCUT&Tag detected 31,611 and 38,992 peaks, respectively. Among them, 17,738 (∼56% of 0.1-g nCUT&Tag) and 19,585 (∼50% of 0.01-g nCUT&Tag) peaks were detected by the 1-g eChIP-seq (**Figures 5B-E**). The signal showed lower enrichment at TSS than the regular nCUT&Tag and eChIP-seq (**Supplementary Figure 3B**). The peak distribution was also a little bit different from the regular nCUT&Tag and eChIP-seq, with less proportion of first-exon peaks and higher proportion of other-exon peaks (**Supplementary Figure 4** and **Supplementary Table 2**). Overall, the low-input nCUT&Tag mapped many peaks commonly as detected by regular nCUT&Tag and eChIP-seq, but the signal was lower and the peak distribution was different. This means that it needs to be improved in the further.

nCUT&Tag Is Scalable for Chromatin Profiling in Other Plant Species

Enhanced ChIP-seq has been used to map high-quality reference epigenomes in rice and *B. napus* (Zhang et al., 2020; Zhao et al., 2020). To examine whether nCUT&Tag could be applied to other plant species, we generated the nCUT&Tag data of the H3K4me3 antibody with crosslinked leaves of the dicot rapeseed (*B. napus*) (**Figure 6A**). The rapeseed nCUT&Tag data showed a high correlation with the eChIP-seq (r = 0.92, Spearman's correlation) (**Figure 6B** and **Supplementary Figure 1D**). The rapeseed nCUT&Tag totally identified 42,984 peaks and there were 10,868 peaks (∼20%) detected by eChIP-seq only (**Figure 6C**). However, there was slight signal enrichment in nCUT&Tag libraries at the 10,868 eChIP-seq unique peak regions (**Figure 6D**). As talked about in the rice fresh nCUT&Tag section, this may be due to the lower sequencing depth relative to the eChIP-seq libraries (**Supplementary Table 4**). The rapeseed nCUT&Tag showed similar signal profiles and peak distribution profiles (**Supplementary Figures 3C**, **4**). The results indicate that nCUT&Tag can be used to study the chromatin landscapes in both monocots and dicots.

DISCUSSION

The standard ChIP-seq (Kaufmann et al., 2010) and eChIP-seq (Zhao et al., 2020) protocols for plants start with fresh tissue, followed by crosslinking, nuclei isolation, sonication, immunoprecipitation, reverse crosslinking, DNA extraction, and library preparation (including end repair, A-tailing, adaptor ligation, and PCR enrichment) (**Table 1**). The procedures are quite complex and require significant input samples and much time for completion. By contrast, nCUT&Tag is a crosslinking-free, sonication-free, immunoprecipitation-free strategy for *in situ* and *in vivo* detection of protein–DNA interactions (**Table 1**). It is a rapid and efficient protocol that all the procedures can be finished within 1 day with as little as 0.01 g of plant tissue.

The sonication-based ChIP-seq assays might underpresent weak or indirect protein–DNA interactions, which might be disrupted during sonication (Fullwood and Ruan, 2009; Bi et al., 2017). For instance, the *Arabidopsis* NUP1 is a nuclear periphery-located protein that loosely interacts with repressive chromatin (Bi et al., 2017). With regular ChIP-seq procedures, the NUP1 peak signals cannot be detected. However, the RE-ChIP-seq (restriction enzyme-mediated ChIP-seq), in which the sonication-based chromatin fragmentation is replaced with restriction enzyme digestion, causes less disruption to protein–DNA interactions and observes signal enrichment of the loosely interacted chromatin positioned around the nuclear periphery (Bi et al., 2017). nCUT&Tag is a sonication-free method and detects ∼6000 unique peaks compared to eChIP-seq (**Figures 3A,C-D**). These peaks show

TABLE 1 | Comparison of nCUT&Tag with our previously reported eChIP-seq protocol.

	nCUT&Tag	eChIP-seq
Input samples	Fixed/fresh, low-input, scalable for single cells	Fixed, low-input
Sonication	No, sonication-free	Yes
Immunoprecipitation	No, immunoprecipitation-free	Protein G beads; low salt; high salt; LiCl buffer
Reverse crosslinking	Direct DNA purification for fresh tissue	Yes
Library preparation	PCR directly	End repair; A-tailing; adaptor ligation; PCR enrichment
Time from tissue to library	1 day	4 days

narrower and weaker signals than those of commonly detected by nCUT&Tag and eChIP-seq (**Figure 3D**), suggesting that they are weak modification sites that are not efficiently preserved during sonication and thus cannot be detected in ChIP-seq assays. Therefore, the *in situ* method nCUT&Tag may have a broader spectrum in mapping *in vivo* protein–DNA interactions, especially for the weak or indirect interactions.

It is a key aspect of epigenomic study to map global chromatin features for understanding transcriptional regulation at single-cell levels. Currently, it is not realistic to perform sonication for a single cell. Therefore, the regular sonication-based ChIP-seq protocols are not suitable for single-cell epigenomic study. However, the PAT- or PGT-mediated sonication-free strategies such as CUT&Tag, ACT-seq, and CoBATCH can be used for single-cell, as well as high-throughput chromatin profiling (Carter et al., 2019; Kaya-Okur et al., 2019; Wang et al., 2019). Hence, nCUT&Tag may be scalable for high-throughput or single-nucleus profiling of histone marks in plants. Importantly, the PAT- or PGT-mediated chromatin immunocleavage strategies may greatly facilitate the development of single-cell ligation-free 3D genome mapping technologies (Ouyang et al., 2020).

Most recently, Liu et al. (2020) developed small-scale Tn5-assisted chromatin cleavage with sequencing (Stacc-seq) to map genome-wide occupancy of RNA polymerase II. The principle of Stacc-seq is similar to CUT&Tag, but the procedures are different. Stacc-seq starts with *in vitro* pre-incubation of antibody with PAT/PGT, followed by incubation of antibody-PAT/PGT complex with live cells (Liu et al., 2020). Compared to CUT&Tag, Stacc-seq adopts only one round of *in vivo* incubation, omitting many buffer-exchange steps. Hence, Stacc-seq can be used rapid profiling of histone marks and transcriptional factor occupancies with hundreds of cells. We believe that Stacc-seq, as well as nCUT&Tag, will be useful alternative methods of ChIP-seq.

CONCLUSION

nCUT&Tag is a simple, rapid, and efficient method that is versatile for studying both active and repressive histone modifications across fresh and crosslinked plant tissues. It is a sonication-free and immunoprecipitation-free protocol that is scalable for single-nucleus chromatin profiling. Moreover, all the procedures in nCUT&Tag can be performed within 1 day with considerably low-input samples, paving a new avenue for rapid single-cell epigenomic studies in plants.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA671638.

AUTHOR CONTRIBUTIONS

WO and XL designed the experiments and wrote the manuscript. XL supervised the research. WO generated data with assistance from YP, QZ, and ZC. XZ, WO, and GL performed data analysis. All authors participated in data interpretation.

FUNDING

This work was supported by the National Natural Science Foundation of China (32070612 and 31771422), the National Key Research and Development Program of China (2016YFD0100904), and the open funds of the National Key Laboratory of Crop Genetic Improvement (ZK201906).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.634679/full#supplementary-material

Supplementary Figure 1 | Correlation heatmaps for the nCUT&Tag and eChIP-seq data.

Supplementary Figure 2 | Scatter plots for the nCUT&Tag and eChIP-seq data.

Supplementary Figure 3 | Heatmaps for nCUT&Tag and eChIP-seq signals across gene body.

Supplementary Figure 4 | Annotation of peaks for the H3K4me3 and H3K9me2 histone marks from nCUT&Tag and eChIP-seq data.

Supplementary Figure 5 | Comparison of the chromatin landscape mapped by fresh nCUT&Tag and crosslinked nCUT&Tag.

Supplementary Table 1 | Summary of the rice nCUT&Tag and eChIP-seq data.

Supplementary Table 2 | Summary of the peak annotation.

Supplementary Table 3 | Summary of peak numbers and FRiP values under various sequencing depths.

Supplementary Table 4 | Summary of nCUT&Tag and eChIP-seq data from the *Brassica napus* young leaves.

REFERENCES

Ai, S., Xiong, H., Li, C. C., Luo, Y., Shi, Q., Liu, Y., et al. (2019). Profiling chromatin states using single-cell itChIP-seq. *Nat. Cell Biol.* 21, 1164–1172. doi: 10.1038/s41556-019-0383-5

Bi, X., Cheng, Y. J., Hu, B., Ma, X., Wu, R., Wang, J. W., et al. (2017). Nonrandom domain organization of the *Arabidopsis* genome at the nuclear periphery. *Genome Res.* 27, 1162–1173. doi: 10.1101/gr.215186.116

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

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218. doi: 10.1038/nmeth.2688

Carter, B., Ku, W. L., Kang, J. Y., Hu, G., Perrie, J., Tang, Q., et al. (2019). Mapping histone modifications in low cell number and single cells using antibody-guided chromatin tagmentation (ACT-seq). *Nat. Commun.* 10:3747. doi: 10.1038/s41467-019-11559-1

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

Fullwood, M. J., and Ruan, Y. (2009). ChIP-based methods for the identification of long-range chromatin interactions. *J. Cell. Biochem.* 107, 30–39. doi: 10.1002/jcb.22116

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004

Johnson, D. S., Mortazavi, A., Myers, R. M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316, 1497–1502. doi: 10.1126/science.1141319

Kaufmann, K., Muino, J. M., Osteras, M., Farinelli, L., Krajewski, P., and Angenent, G. C. (2010). Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nat. Protoc.* 5, 457–472. doi: 10.1038/nprot.2009.244

Kaya-Okur, H. S., Wu, S. J., Codomo, C. A., Pledger, E. S., Bryson, T. D., Henikoff, J. G., et al. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* 10:1930. doi: 10.1038/s41467-019-09982-5

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Liu, B., Xu, Q., Wang, Q., Feng, S., Lai, F., Wang, P., et al. (2020). The landscape of RNA Pol II binding reveals a stepwise transition during ZGA. *Nature* 587, 139–144. doi: 10.1038/s41586-020-2847-y

Ouyang, W., Xiong, D., Li, G., and Li, X. (2020). Unraveling the 3D genome architecture in plants: present and future. *Mol. Plant* 13, 1676–1693. doi: 10.1016/j.molp.2020.10.002

Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42, W187–W191. doi: 10.1093/nar/gku365

Skene, P. J., Henikoff, J. G., and Henikoff, S. (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* 13, 1006–1019. doi: 10.1038/nprot.2018.015

Sun, Y., Dong, L., Zhang, Y., Lin, D., Xu, W., Ke, C., et al. (2020). 3D genome architecture coordinates trans and cis regulation of differentially expressed ear and tassel genes in maize. *Genome Biol.* 21:143. doi: 10.1186/s13059-020-02063-7

Tao, X., Feng, S., Zhao, T., and Guan, X. (2020). Efficient chromatin profiling of H3K4me3 modification in cotton using CUT&Tag. *Plant Methods* 16:120. doi: 10.1186/s13007-020-00664-8

Wang, Q., Xiong, H., Ai, S., Yu, X., Liu, Y., Zhang, J., et al. (2019). CoBATCH for high-throughput single-cell epigenomic profiling. *Mol. Cell* 76, 206.e7–216.e7. doi: 10.1016/j.molcel.2019.07.015

Zhang, J., Chen, L. L., Xing, F., Kudrna, D. A., Yao, W., Copetti, D., et al. (2016). Extensive sequence divergence between the reference genomes of two elite indica rice varieties Zhenshan 97 and Minghui 63. *Proc. Natl. Acad. Sci. U.S.A.* 113, E5163–E5171. doi: 10.1073/pnas.1611012113

Zhang, Q., Guan, P., Zhao, L., Ma, M., Xie, L., Yue, L., et al. (2020). Asymmetric epigenome maps of subgenomes reveal imbalanced transcription and distinct evolutionary trends in *Brassica napus*. *Mol. Plant* [Epub ahead of print] doi: 10.1016/j.molp.2020.12.020

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137

Zhao, L., Xie, L., Zhang, Q., Ouyang, W., Deng, L., Guan, P., et al. (2020). Integrative analysis of reference epigenomes in 20 rice varieties. *Nat. Commun.* 11:2658. doi: 10.1038/s41467-020-16457-5

Zheng, X. Y., and Gehring, M. (2019). Low-input chromatin profiling in *Arabidopsis* endosperm using CUT&RUN. *Plant Reprod.* 32, 63–75. doi: 10.1007/s00497-018-00358-1

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.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 June 2021
doi: 10.3389/fpls.2021.685189

Transcriptome-Wide Analysis of RNA m^{6}A Methylation and Gene Expression Changes Among Two *Arabidopsis* Ecotypes and Their Reciprocal Hybrids

*Zhihui Xu[1,2], Xiaobo Shi[2], Mengmei Bao[2], Xiaoqian Song[2], Yuxia Zhang[2], Haiyan Wang[2], Hairong Xie[2], Fei Mao[2], Shuai Wang[2], Hongmei Jin[3], Suomeng Dong[4], Feng Zhang[4], Zhe Wu[5] and Yufeng Wu[2]**

[1] College of Life Science, Nanjing Agricultural University, Nanjing, China, [2] State Key Laboratory for Crop Genetics and Germplasm Enhancement, Jiangsu Key Laboratory for Information Agriculture, Bioinformatics Center, Academy for Advanced Interdisciplinary Studies, Nanjing Agricultural University, Nanjing, China, [3] Institute of Agricultural Resources and Environment, Jiangsu Academy of Agricultural Sciences, Nanjing, China, [4] College of Plant Protection, Nanjing Agricultural University, Nanjing, China, [5] Department of Biology, SUSTech-PKU Institute of Plant and Food Science, Southern University of Science and Technology, Shenzhen, China

The remodeling of transcriptome, epigenome, proteome, and metabolome in hybrids plays an important role in heterosis. N(6)-methyladenosine (m^{6}A) methylation is the most abundant type of post-transcriptional modification for mRNAs, but the pattern of inheritance from parents to hybrids and potential impact on heterosis are largely unknown. We constructed transcriptome-wide mRNA m^{6}A methylation maps of *Arabidopsis thaliana* Col-0 and *Landsberg erecta* (Ler) and their reciprocal F_1 hybrids. Generally, the transcriptome-wide pattern of m^{6}A methylation tends to be conserved between accessions. Approximately 74% of m^{6}A methylation peaks are consistent between the parents and hybrids, indicating that a majority of the m^{6}A methylation is maintained after hybridization. We found a significant association between differential expression and differential m^{6}A modification, and between non-additive expression and non-additive methylation on the same gene. The overall RNA m^{6}A level between Col-0 and Ler is clearly different but tended to disappear at the allelic sites in the hybrids. Interestingly, many enriched biological functions of genes with differential m^{6}A modification between parents and hybrids are also conserved, including many heterosis-related genes involved in biosynthetic processes of starch. Collectively, our study revealed the overall pattern of inheritance of mRNA m^{6}A modifications from parents to hybrids and a potential new layer of regulatory mechanisms related to heterosis formation.

Keywords: RNA m^{6}A methylation, hybrid, heterosis, *Arabidopsis*, RNA modification dynamics

OPEN ACCESS

Edited by:
Dazhong Dave Zhao, University of Wisconsin–Milwaukee, United States

Reviewed by:
Chuang Ma, Northwest a and F University, China
Hua Jiang, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany

***Correspondence:**
Yufeng Wu
yfwu@njau.edu.cn

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

Received: *24 March 2021*
Accepted: *14 May 2021*
Published: *10 June 2021*

Citation:
Xu Z, Shi X, Bao M, Song X, Zhang Y, Wang H, Xie H, Mao F, Wang S, Jin H, Dong S, Zhang F, Wu Z and Wu Y (2021) Transcriptome-Wide Analysis of RNA m^{6}A Methylation and Gene Expression Changes Among Two Arabidopsis Ecotypes and Their Reciprocal Hybrids. Front. Plant Sci. 12:685189. doi: 10.3389/fpls.2021.685189

SIGNIFICANCE STATEMENT

The reprogramming and corresponding effect of mRNA m^{6}A methylation on hybrids remain highly unknown. We demonstrated the pattern of conserved inheritance of m^{6}A methylation from parents to hybrids and the potential impact on heterosis formation, uncovering mRNA m^{6}A methylation as a new layer of regulatory mechanisms in the formation of hybrid vigor.

INTRODUCTION

Heterosis refers to the increased performance of hybrid offspring relative to their parents in many traits, such as growth rate and biomass (Birchler et al., 2003, 2010; Hochholdinger and Hoecker, 2007; Chen, 2010; Birchler, 2015). Both genetic and epigenetic mechanisms are thought to be involved in heterosis (Chen, 2013). Epigenetic changes have been found to impact hybrid vigor (Cubas et al., 1999; Manning et al., 2006; Shindo et al., 2006; Ni et al., 2009; He et al., 2010). DNA methylation level is altered by trans-chromosomal methylation (TCM) and trans-chromosomal demethylation (TCdM) (Greaves et al., 2014), which changes the overall DNA methylation level in the F_1 hybrids, especially in regions that are differentially methylated in two parents (Shen et al., 2012). Histone modification patterns in hybrids of rice or maize have shown correlations between altered gene expression and changes in histone marks compared with the parents (He et al., 2010, 2013; Lv et al., 2019). In *Arabidopsis* hybrids, global histone modifications of the parents are largely transmitted to the F_1 generation (Moghaddam et al., 2011; Dong et al., 2012; Yang et al., 2016). DNA methylation and histone modifications are altered at many loci, such as circadian clock associated1 (CCA1) and late elongated hypocotyl (LHY), which are associated with growth vigor in *Arabidopsis* F_1 hybrids (Ni et al., 2009; Shen et al., 2012).

Recently, chemical modifications of mRNAs, such as N(6)-methyladenosine (m^{6}A), N(1)-methyladenosine (m^{1}A), and 5-methylcytosine (m^{5}C), have emerged as an additional level of transcript regulation (Dominissini et al., 2012, 2016; Meyer et al., 2012; Li et al., 2016, 2017). m^{6}A methylation is the most abundant type of modification for mRNAs, occurring in more than one-third of mammalian transcripts and half of the plant transcripts (Dominissini et al., 2012; Meyer et al., 2012; Li et al., 2014; Luo et al., 2014; Luo et al., 2020; Wan et al., 2015; Zhou et al., 2019; Miao et al., 2020). The m^{6}A modification is reversible and dynamic, with m^{6}A demethylase acting as an eraser and methyltransferase acting as a writer (Jia et al., 2011; Meyer and Jaffrey, 2017). Recognition of these dynamic m^{6}A modifications by YTH domain-containing proteins leads to a broad range of functions associated with the change in mRNA stability, cap-independent translation, splicing, translation efficiency, and mRNA structure (Dominissini et al., 2012; Meyer et al., 2015; Meyer and Jaffrey, 2017), but the location of m^{6}A in mRNA determines different functions (Gilbert et al., 2016). In the 5' UTR, m^{6}A participates in mRNA cap-independent translation by directly binding to eukaryotic initiation factor 3 (eIF3) and then recruiting the 40S ribosomal subunit to initiate translation (Meyer et al., 2015). On the other hand, m^{6}A in the 3' UTR has been reported to have several functions, such as promoting translation by binding with METTL3 and eIF3h to facilitate formation of the translation loop (Choe et al., 2018), regulating mRNA lifetime by binding with YTHDF2, which relocates transcripts to the P-body (Wang et al., 2014), and changing mRNA structure to affect RNA-protein interactions (Liu et al., 2015).

In *Arabidopsis thaliana*, m^{6}A is essential in embryo development (Zhong et al., 2008). Further research revealed that m^{6}A is also essential in post-embryonic development (Bodi et al., 2012), for example, for normal trichome morphology and correct timing of leaf formation (Arribas-Hernandez et al., 2018; Scutenaire et al., 2018; Wei et al., 2018), partly because it regulates the expression of key shoot meristem genes to control shoot apical meristem (SAM) proliferation (Shen et al., 2016). Transcriptome-wide mapping of m^{6}A in *Arabidopsis* wild-type (WT) and related mutants indicated a complex relationship between m^{6}A modifications and gene expression. Lack of FKBP12 interacting protein 37 (FIP37), a component of the methyltransferase complex in *Arabidopsis,* results in a dramatically reduced abundance of m^{6}A, as most transcripts bearing m^{6}A in WT are decreased in the mutant (Shen et al., 2016). In addition, further study showed that m^{6}A inhibits mRNA degradation through inhibition of site-specific cleavage (Anderson et al., 2018). Nevertheless, it was reported that the highly expressed transcripts had fewer m^{6}A modifications, as revealed by transcriptome-wide m^{6}A patterns in *Arabidopsis* (Wan et al., 2015). Although m^{6}A abundance varies among *Arabidopsis* accessions and affects transcript abundance, how m^{6}A changes in F_1 hybrids relative to their parents and its potential role in determining F_1 hybrid vigor have not been clarified.

In this study, we selected two *Arabidopsis* ecotypes, namely, Col-0 and *Landsberg erecta* (Ler), and their F_1 reciprocal hybrids, to investigate the potential effect of m^{6}A on heterosis. We identified the distribution pattern and the intensity change in m^{6}A in Col-0, Ler, and their F_1 reciprocal hybrids. We showed that the peaks and distribution features of m^{6}A methylation are highly conserved between accessions. Although changes in m^{6}A intensity and transcript abundance within accessions are weakly positively correlated, upregulation of m^{6}A between accessions tends to be associated with a downregulated abundance of mRNA and vice versa. We found that the overall m^{6}A difference between the parents is attenuated at allelic sites in the hybrids, and that there is a negative correlation between the expression and corresponding m^{6}A intensity of allelic genes. Interestingly, even though hundreds of m^{6}A peaks are changed between the parents and hybrids, many biological functions of the corresponding genes are consistently affected, including the biosynthetic processes of starch, which have been reported to be associated with growth vigor. The data, therefore, suggest the overall pattern of mRNA m^{6}A remodeling in hybrids, which may contribute to heterosis formation.

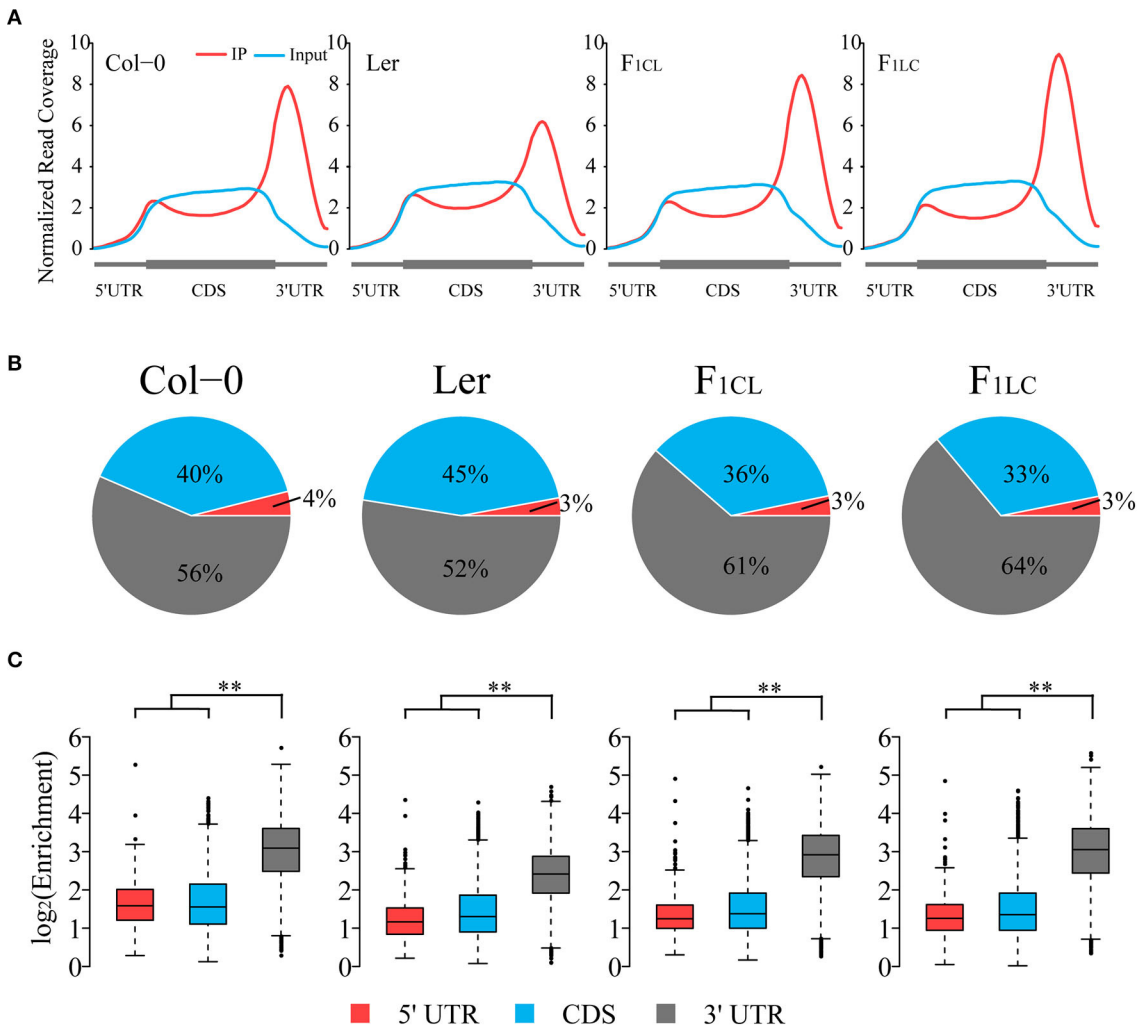

FIGURE 1 | Global pattern of m^{6}A peaks in Col-0, Ler, and their F_1 reciprocal hybrids. **(A)** Coverage of normalized reads along transcripts. Each transcript is divided into three non-overlapping features: 5' UTR, CDS, and 3' UTR. **(B)** Distribution of m^{6}A peaks in transcript features of parents and hybrids. **(C)** Relative enrichment of m^{6}A peaks of each transcript feature. Enrichment = Normalized m^{6}A-seq reads divided by normalized input reads of each peak. $^{**}p < 2.2e{-}16$, Wilcoxon rank-sum test.

RESULTS

Transcriptome-Wide Profile of m^{6}A Methylation Among Col-0, Ler, and Their F_1 Reciprocal Hybrids

To explore RNA m^{6}A abundance variation between the two ecotypes and its alteration in hybrids, we first analyzed transcriptome-wide m^{6}A profiles among Col-0, Ler, and their F_1 reciprocal hybrids (**Supplementary Figure 1A**) by applying N^6-methyladenosine sequencing (m^{6}A-seq) with two biological replicates. Sequencing data of RNA input and immunoprecipitation (RIP) are highly correlated between replicates, indicating the high quality of m^{6}A-seq in this study (**Supplementary Figures 1B,C**). We found that the normalized reads from m^{6}A-RIP of all samples are enriched in the 3' UTR of the transcripts (**Figure 1A**), which is similar to the results of previous research (Luo et al., 2014; Wan et al., 2015). The normalized read depth in Ler is significantly lower than that of the other three samples, suggesting that the overall m^{6}A abundance of Ler was lower (**Figure 1A**). To exclude the possible bias introduced by the reference genome, we performed exact analysis using the Ler reference genome rather than Col-0 and still obtained identical results (**Supplementary Figure 2**). Interestingly, we did not find low m^{6}A abundance in the 3' UTR of the two hybrids, similar to Ler.

To further study global patterns of m^{6}A in Col-0, Ler and their hybrids, we identified m^{6}A peaks using a transcriptomic peak caller, METPeak (Cui et al., 2016). A total of 13,145, 13,562, 12,956, and 12,542 peaks are detected in Col-0, Ler, F_{1CL}, and F_{1LC}, respectively (**Supplementary Table 1**); and these peaks were located in ~9,778, 9,920, 10,066, and 10,017 protein-coding

genes, respectively. The majority of these genes have one or two m^{6}A sites (**Supplementary Figure 3A**), which is consistent with a previous report (Wan et al., 2015). In agreement with the distribution of m^{6}A-seq reads, the majority of the m^{6}A peaks are enriched in the 3' UTR and CDS region, while only 3–4% of the m^{6}A peaks are located in the 5' UTR (**Figure 1B**). The enrichment degree of peaks in the 3' UTR is significantly higher (Wilcoxon rank sum test, $p < 2.2e{-}16$) than that of peaks in the 5' UTR and CDS among the four samples (**Figure 1C**). As expected, we also found that the enrichment of the m^{6}A peaks in Ler is significantly lower than that in the other groups (Wilcoxon rank sum test, $p < 2.2e{-}16$, **Supplementary Figure 3B**).

To further analyze the feature of the distribution of m^{6}A peaks, we counted the number of peaks around the start codon segment and the stop codon segment (200 nt centered on the start codon and stop codon, respectively), and found that ∼40% of the peaks are located in these two regions (**Supplementary Figure 3C**). The number of peaks in the start codon is relatively low in all four samples. However, there are more than 4,000 m^{6}A peaks located in the stop codon segment (**Supplementary Figure 3C**), which is consistent with previous findings in mammals and plants showing that m^{6}A peaks are preferentially located around stop codons (Dominissini et al., 2012; Luo et al., 2014; Wan et al., 2015).

Variations of m^{6}A Modification Among the Parental Lines and Hybrids

Previous research has shown that m^{6}A is highly conserved between two accessions of *Arabidopsis*, namely, Can-0 and Hen-16 (Luo et al., 2014). We found that 10,584 m^{6}A peaks (80.5% of Col-0, 78% of Ler; **Figure 2A**) are common between Col-0 and Ler, and that these peaks are located in 8,302 expressed transcripts (49.4% of the total). In addition, we found that the majority of the m^{6}A peaks are common among the parental lines and F_1 hybrids. There are 9,641 (74.4% of F_{1CL}) and 9,331 (74.4% of F_{1LC}) m^{6}A peaks that are common between the parents and the F_1 hybrids, respectively (**Figure 2B**). These peaks are also located in 7,844 and 7,723 of the expressed transcripts in F_{1CL} and F_{1LC}, respectively (**Supplementary Figure 3E**). The common peaks (11,000) between F_{1CL} and F_{1LC} account for 85.6–88.5% of the total peaks in F_1 hybrids (**Figure 2A**), which is slightly higher than that in the two parents. Collectively, these data indicate a more general conservation pattern of RNA m^{6}A modification among accessions and hybrids of *Arabidopsis*.

Considering the obvious difference in m^{6}A levels between Col-0 and Ler, it is necessary to determine whether common m^{6}A peaks between any two samples are significant differentially methylated peaks (DMPs). We established two criteria for DMPs: (1) passed Fisher's exact test after multiple comparison corrections (FDR < 0.05); (2) the difference in peak enrichment between any two samples was larger than a 1.5-fold change. Eventually, we identified 1,776 DMPs (16.8% of the common peaks) between the parents, among which the intensity of 1,721 (16.3%) peaks, as expected, is higher in Col-0 (**Figures 2C,D**, **Supplementary Table 2**). For the comparison between F_1 reciprocal hybrids, we found only 2 DMPs (0.02%), suggesting that paternal or maternal effects on the level of m^{6}A modifications are weak in *Arabidopsis* (**Figure 2C**). For the m^{6}A peaks shared between the parents and F_{1CL} or between the parents and F_{1LC}, we identified 315, 479, 477, and 1,273 DMPs, respectively (**Figure 2C**). Taken together, the intensity of common m^{6}A peaks tends to be conserved between accessions or during inheritance from parents to hybrids.

Relationship Between Transcript Abundance and m^{6}A Modification Level

Multiple recent studies have indicated complex functions of m^{6}A in transcription regulation with the ability to stabilize (Luo et al., 2014; Anderson et al., 2018) or destabilize mRNAs in *Arabidopsis* (Wan et al., 2015). We analyzed the relationships of transcript abundance and the corresponding m^{6}A levels. We found a weak positive correlation between the expression abundance and intensity of m^{6}A modification on one gene within each accession (**Figure 3A**). Overall, the genes with m^{6}A modification show significantly higher expression than non-m^{6}A-containing genes (**Figure 3B**). In addition, more than 60% of the expressed genes are associated with at least one m^{6}A peak (**Supplementary Figure 3A**, **Supplementary Table 1**).

Next, we investigated the relationship between changes in m^{6}A methylation and transcript abundance in the parent lines and their F_1 reciprocal hybrids. We first identified differentially expressed genes (DEGs) between the lines (**Supplementary Figures 4A,B**) and checked the overlap between DEGs and DMPs. We found that the proportion of DEGs associated with DMPs is significantly higher than that of non-DEGs (**Figure 3C**). Even so, only 3.29–13.26% of the DEGs are associated with DMPs between the parent lines and hybrids (**Figure 3C**). Taken together, these results indicated that changes in m^{6}A intensity on transcripts tend to be associated with changes in abundance, and that most DEGs are not directly associated with m^{6}A changes in *Arabidopsis*.

We then focused on genes with significant changes in both expression and m^{6}A modification between accessions. Most DMPs showed upregulated m^{6}A intensity in Col-0 between Col-0/Ler (comparison between Col-0 and Ler), as well as between Col-0/F_{1CL} and between Col-0/F_{1LC} (**Figure 3D**). A total of 862 DEGs between Col-0/Ler are associated with DMPs upregulated in Col-0, among which there are significantly more downregulated expressed genes than upregulated genes in Col-0 (**Figure 3D**, $p = 2.41e{-}5$, chi-square test). A similar pattern is also found in Col-0/F_{1CL} ($p = 1.7e{-}10$) and Col-0/F_{1LC} ($p = 0.037$, **Figure 3D**). There are more DMPs showing downregulated m^{6}A intensity in Ler between Ler/F_{1CL} and between Ler/F_{1LC}, and these DMPs are also associated with more genes with upregulated expression in Ler (**Supplementary Figure 4C**, $p = 1.19e{-}12$ for Ler/F_{1CL}; $p = 2.38e{-}20$ for Ler/F_{1LC}). This result indicates that downregulated DMPs tend to be associated with more upregulated DEGs and vice

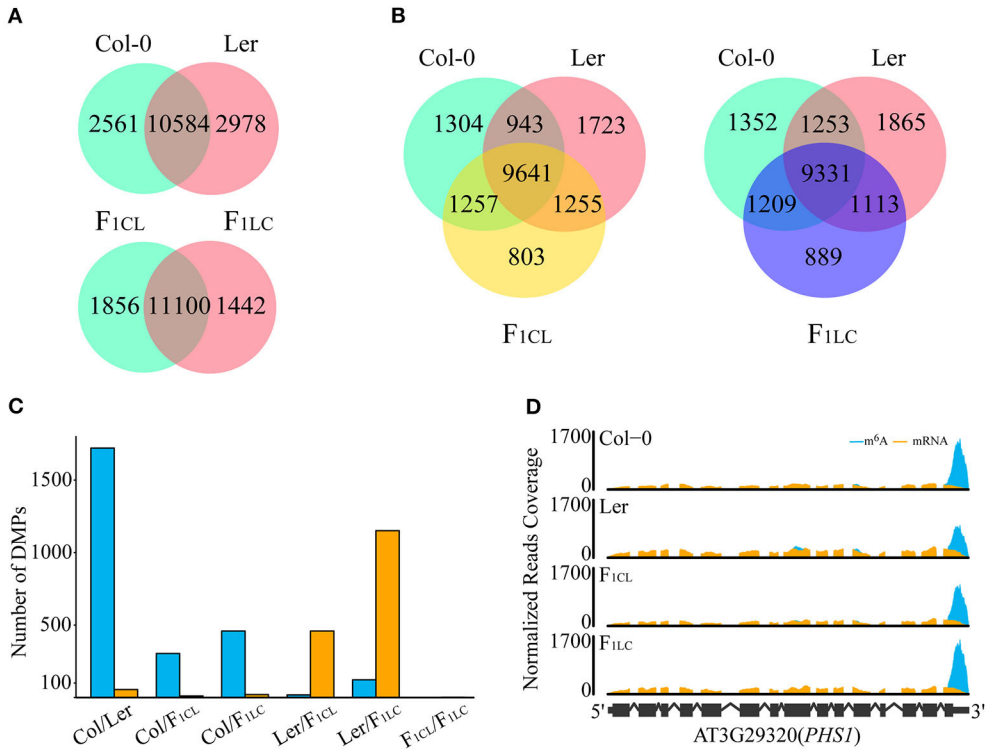

FIGURE 2 | Differences in m^{6}A modifications among Col-0, Ler, and their F_1 reciprocal hybrids. **(A,B)** Number of shared m^{6}A peaks between accessions. **(C)** Number of DMPs in each comparison. Blue bars, DMPs showing upregulated m^{6}A intensity in the former comparison. Orange bars, DMPs showing upregulated m^{6}A intensity in the latter. **(D)** Diagram for differentially methylated peaks. m^{6}A, normalized IP reads; mRNA, normalized input reads.

versa between accessions of *Arabidopsis*, implying that the complexity of m^{6}A function affects the transcript abundance of genes.

Relationship Between Non-additive Expression and Non-additive m^{6}A Modification

We identified 2,758 and 4,123 genes showing non-additive expression in F_{1CL} and F_{1LC}, respectively. Similar to gene expression, the inheritance of m^{6}A modifications in hybrids can be additive or non-additive. We defined m^{6}A peaks with a significant change between enrichment value in hybrid and the average enrichment value of parents (MPV) (FDR $<$ 0.05, see methods for detail) as non-additive m^{6}A modified peaks. The majority (95.6 and 95.2%) of the m^{6}A peaks show additive patterns in both hybrids, while only 538 and 563 peaks in F_{1CL} and F_{1LC}, respectively, are non-additive (**Supplementary Table 3**). Moreover, non-additive m^{6}A peaks are significantly associated with non-additively expressed genes in both hybrids (**Supplementary Table 3**, $p < 2.2e{-}16$ for both F_{1CL} and F_{1LC}, chi-square test). We still observed that only 6.53–6.82% of non-additively expressed genes show a non-additive pattern of m^{6}A modification, indicating that m^{6}A may play a role in the regulation of non-additive gene expression in *Arabidopsis* hybrids.

Relationship Between Allelic Gene Expression and Allelic m^{6}A Methylation in F_1 Hybrids

To analyze the allelic bias in gene expression and m^{6}A modifications in hybrids, we identified single-nucleotide polymorphisms (SNPs) between Col-0 and Ler with stringent criteria (see methods) and used these SNPs to determine the reads of RNA-seq or m^{6}A-seq generated from the allele of Col-0 or Ler. A total of 76,983 SNPs with high confidence are identified. These SNPs associate with 8,972 and 8,991 genes and with 2,509 and 2,325 m^{6}A peaks in F_{1CL} and F_{1LC}, respectively, which are used in the following analysis. As expected, we still observed significantly higher m^{6}A modification in Col-0 than in Ler on these SNPs (**Figure 4A**). Nevertheless, this bias tends to disappear between the two parental alleles in the hybrids. The log-transformed mean value of the m^{6}A ratio between the two allelic SNPs is close to zero, and the majority of the ratio (94.1% for F_{1CL} and 93.6% for F_{1LC}) falls within the interval $(-1, 1)$ in both hybrids (**Figure 4A**), indicating that the overall m^{6}A difference between the parents is attenuated at allelic sites in the hybrids. The pattern of attenuation is not observed for the expression of allelic genes (**Figure 4B**). We have identified only

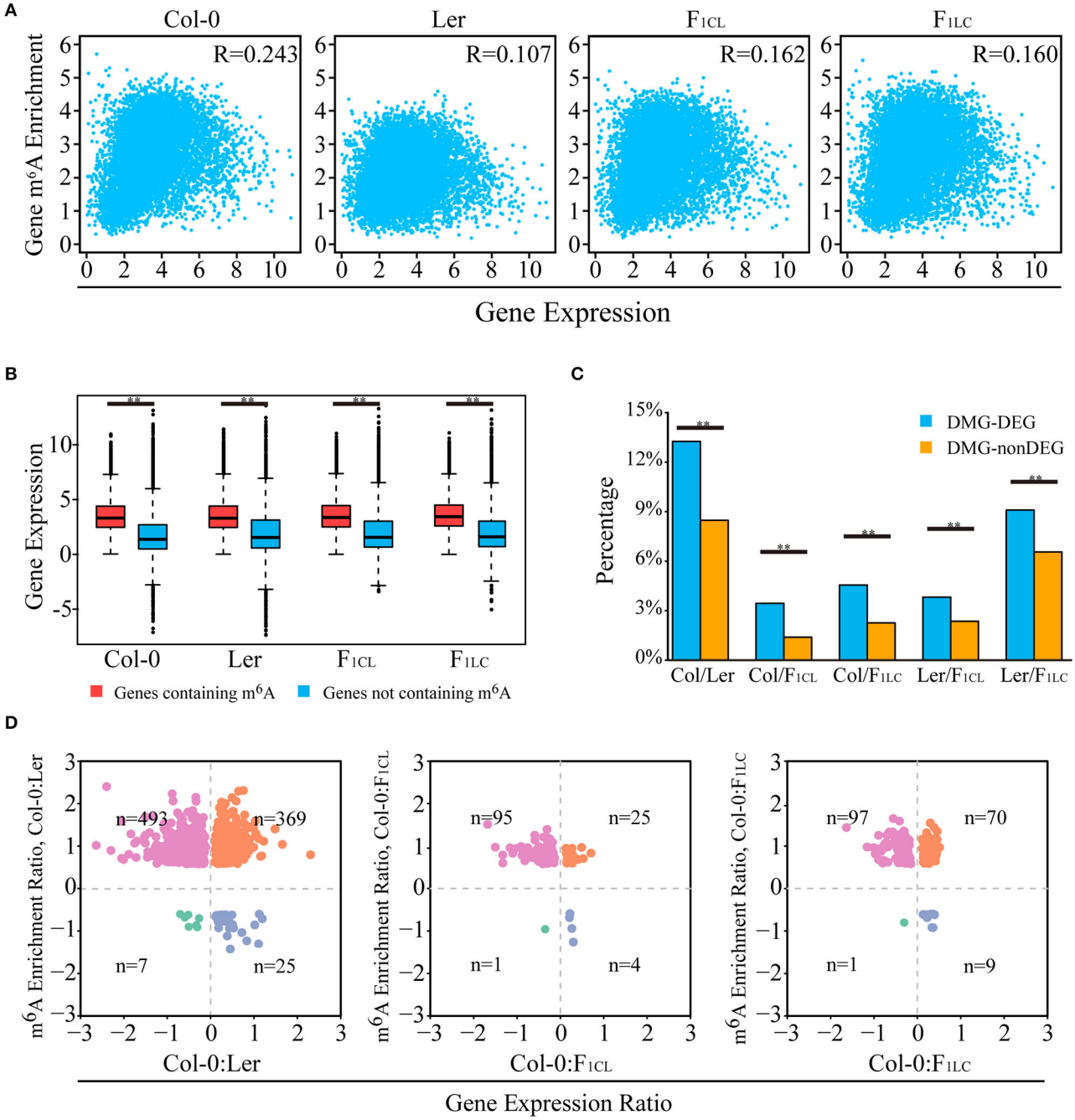

FIGURE 3 | Relationship between m^{6}A methylation level and transcript abundance. **(A)** Scatter plot showing the correlation of m^{6}A modification and transcript abundance. R indicates Pearson correlation coefficient. **(B)** Transcripts with m^{6}A peaks showing higher abundance levels. **(C)** Percentage of DEGs (differentially expressed genes) overlapping with DMGs (genes with differentially m^{6}A-methylated peaks), indicated as DMG-DEGs, and percentage of non-DEGs associated with DMGs, indicated as DMG-nonDEGs. $^{**}p < 1e{-}6$, Wilcoxon rank-sum test. **(D)** Scatter plot of DMG-DEGs between accessions showing the relationship of m^{6}A modification and transcript abundance. For example, the m^{6}A enrichment ratio of Col-0: Ler is calculated as log2 (enrichment of Col-0/enrichment of Ler) of m^{6}A peaks. The gene expression ratio of Col-0: Ler is calculated as log2 (FPKM of Col-0/FPKM of Ler) of transcripts. n indicates number of DMG-DEGs in each quadrant. For **(A–D)**, gene m^{6}A enrichment is calculated by normalized m^{6}A-seq reads number divided by normalized input reads of peaks within the transcript, and gene expression is indicated by the FPKM of the input RNA-seq data.

four and seven peaks showing significant allele-specific RNA m^{6}A methylation (FDR < 0.05, see Methods) in the hybrids, implying extremely rare allele bias of RNA m^{6}A methylation after the combination of the two parental genomes. Despite the smaller difference in m^{6}A abundance between the alleles, the correlation between the allelic abundance of mRNA and the

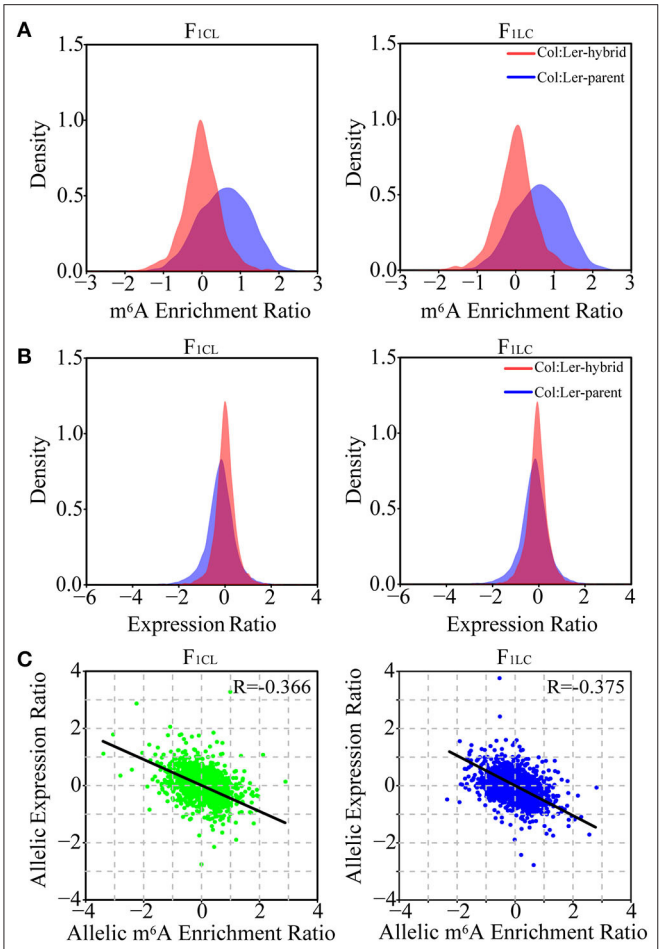

FIGURE 4 | Relationship between allelic gene expression and allelic m^{6}A methylation in F_1 hybrids. Density distribution of the **(A)** m^{6}A enrichment ratio and **(B)** gene expression ratio between allelic sites of Col-0 and Ler (marked as Col: Ler-parent) and the ratio between allelic sites inherited from Col-0 and Ler in the hybrids (Col: Ler-hybrid). **(C)** Scatter plot showing the correlation between allele-specific expression and allele-specific m^{6}A methylation in the hybrids. Only the reads of m^{6}A-seq or RNA-seq mapped to the SNPs between Col-0 and Ler with high confidence are used in the analysis.

corresponding allelic intensity of m^{6}A methylation is negative (**Figure 4C**). This result is consistent with the relationship between DEGs and DMPs.

Biological Function of Genes Associated With Significant Changes in m^{6}A

F_1 hybrids crossed by ecotypes of *Arabidopsis*, as well as Col-0×Ler (Groszmann et al., 2014), showed clear growth vigor (**Supplementary Figure 1**), but the relationship between heterosis and changes in m^{6}A abundance between the parent lines and hybrids was unknown. We first focused on the function of genes showing significantly differential m^{6}A methylation (**Supplementary Table 4**), which were referred to as differentially m^{6}A-modified genes (DMGs). We identified 462 enriched GO terms of DMGs between Col-0 and Ler, among which 160–294 (34–63%) are also identified as enriched GO terms of DMGs generated from the comparisons between the parents and hybrids (**Supplementary Table 5**). Interestingly, the enriched GO terms of DMGs between the parents and hybrids tend to be consistent. For instance, we found 319 enriched GO terms of DMGs between Col-0/F_{1CL}, among which 231–267 (72–84%) are also identified in the Col-0/F_{1LC}, Ler/F_{1CL}, and Ler/F_{1LC} comparisons. These data implied that there is clear heterogeneity of biological functions affected by differential m^{6}A modification between Col/Ler (between-parent difference) and between parent/hybrid (parent-hybrid difference). We kept the enriched GO terms of DMGs from the parent-hybrid comparison but not from the between-parent comparison and found a clear trend of enriched biological functions, such as biosynthetic and metabolic processes of multiple carbohydrates, secondary metabolic processes, and development of shoot, root hair, and so on (**Figure 5A**, **Supplementary Table 4**), among which starch biosynthetic process was reported to be involved in heterosis (Chen, 2013).

Hybrid vigor has been found to be related to changes in transcription, epigenetic modifications, and protein abundance (Chen, 2013). Considering that m^{6}A is involved in multiple biological processes related to RNA fate at the post-transcriptional level, it is worthwhile to focus on the function of genes showing differential m^{6}A modification without changes in gene expression. We found that the enriched GO terms of DMGs and not DEGs between parents/hybrids are associated with membrane- or chloroplast-located proteins, transport, or the proteasome complex (**Supplementary Figure 5A**). This pattern is clearly different from the enriched GO terms of the DEGs but not DMGs between parent/hybrid, which are associated with stress response genes, mitochondria-located genes, etc (**Supplementary Figure 5B**). These data implied that m^{6}A modification could be involved in the formation of F_1 hybrid vigor through post-transcriptional regulation of mRNA without changing the abundance.

Several genes involved in starch and carbohydrate metabolism promote growth and biomass vigor in *Arabidopsis* (Chen, 2013), so we focused on the DMGs involved in the starch biosynthetic process (**Figure 5B**). There are 51 DMGs from all four parent/hybrid comparisons annotated as genes of the starch biosynthetic process (GO: 0019252). We checked the published biological functions against the TAIR database one by one and found 20 genes associated with the biomass and growth rate of *Arabidopsis* (**Supplementary Table 6**). We visualized seven genes of 51 DMGs that are annotated as starch metabolism genes through Mapman (Thimm et al., 2004). Interestingly, six genes are located in chloroplasts and involved in the same pathway, and four of them control biomass and growth rate in *Arabidopsis* based on published results (**Figure 5B**). Collectively, these data indicated the strong association between changes in m^{6}A methylation and the growth vigor of F_1 hybrids.

DISCUSSION

Multiple transcriptome-wide maps revealed highly conserved patterns of m^{6}A methylation among *Arabidopsis* accessions

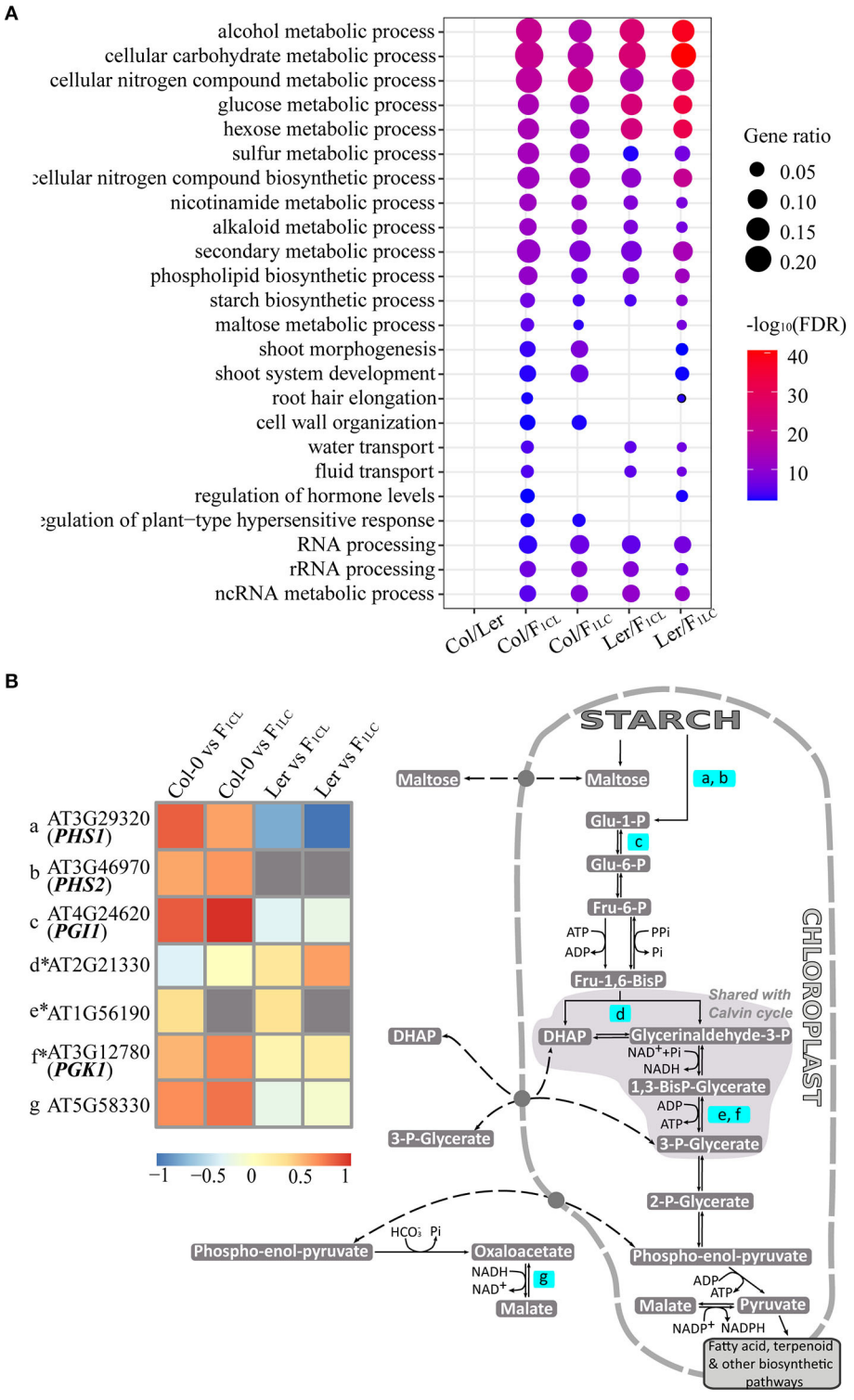

FIGURE 5 | Enriched biological functions of differentially m^{6}A-methylated genes. **(A)** GO terms of genes associated with DMPs that are enriched only in comparisons of parents/hybrids, e.g., m^{6}A-methylated peaks showing significant differences between Col-0 and F_{1CL}. Only some enriched GO terms are shown in the figure, and all GO terms are listed in **Supplementary Table 4**. **(B)** A diagram for genes with DMPs involved in starch biosynthetic process related to biomass in *Arabidopsis*. The pathway is modified based on the Mapman database. Only genes in the pathway with DMPs are highlighted, of which m^{6}A methylation differences between parents/hybrids are shown in the heatmap. *indicates the genes showing non-additive m^{6}A methylation in the hybrids simultaneously.

(Can-0 and Hen-16) or organs (leaf, root, and flower) (Luo et al., 2014; Wan et al., 2015). More than 70% of m^{6}A peaks are shared between *Arabidopsis* Can-0 and Hen-16 (Luo et al., 2014), similar to the percentage (78%) of peaks shared between Col-0 and Ler in this study. We also found that ∼74% of m^{6}A peaks are shared between the parents and hybrids. In addition, our results indicated that m^{6}A modifications in hybrids are enriched around the 3' UTR, stop and start codons of transcripts, showing consistent features across accessions and organs (Luo et al., 2014; Wan et al., 2015). Moreover, RNA m^{6}A methylation peaks are also conserved between two inbred lines (B73 and Mo17) of maize (Luo et al., 2020) and two tissues of rice (Li et al., 2014). In summary, these results implied a more general conservation pattern of m^{6}A methylation in plants, which could be related to the fundamental role of m^{6}A methylation in plant development (Zhong et al., 2008; Shen et al., 2016; Anderson et al., 2018). Nevertheless, a recently published study showed that there are much more genes with differentially m^{6}A level or non-additively m^{6}A variation in maize hybrid (B73 × Mo17) compared with the parents (Luo et al., 2021), implying that the pattern of m^{6}A reprogramming in hybrid is related to species or parent lines with different degree of variation.

The effects of m^{6}A modification on gene expression vary among genes. In an *Arabidopsis* demethylase ALKBH10B loss-of-function mutant, mRNAs of flower development genes, such as FT, SPL3, and SPL9 show increased m^{6}A modification but reduced stability (Duan et al., 2017). Nevertheless, lack of m^{6}A modifications on the mRNA of the WUS and STM genes enhances their stability in the FIP37 mutant line of *Arabidopsis* (Shen et al., 2016). Additional studies have indicated the biological functions of stabilizing or destabilizing mRNAs in *Arabidopsis* (Luo et al., 2014; Wan et al., 2015; Anderson et al., 2018). The data also indicated the conflicting functions of m^{6}A in regulating gene expression. Overall, we observed a very weak positive correlation between the abundance of mRNA and the intensity of m^{6}A modification within each of the accessions. However, we also found that mRNAs with significantly decreased methylation of m^{6}A tend to show upregulated expression between accessions or between parents and hybrids. The complex regulatory roles of m^{6}A in transcript abundance might be correlated with its location (Luo et al., 2014), differences between readers (Wei et al., 2018), or the RNA structure dependent on m^{6}A (Liu et al., 2015).

The molecular mechanism of heterosis is quite complex; and omics methods, ranging from transcriptomics to metabolomics, have provided novel insights into the mechanism (Chen, 2013). Changes in epigenetic modifications, such as histone methylation in hybrids, could promote growth by altering gene expression (Ni et al., 2009). As a newly identified reversible modification of RNA, the reprogramming of m^{6}A in hybrids and the corresponding functions related to heterosis remain largely elusive. The data indicated that most of the differentially expressed genes are not associated with differential m^{6}A methylation, and that only a few hundred m^{6}A peaks are significantly changed between parents and hybrids. However, these peaks are associated with many biological functions, of which 20 of 51 starch- and carbohydrate-related genes are confirmed as being associated with biomass vigor in *Arabidopsis* (**Supplementary Table 6**). We did not identify the genes showing differential m^{6}A methylation involved in the circadian rhythm regulatory network, for instance, LHY, GI, CCA1, and TOC1, which are also related to biomass vigor in hybrids crossed by two accessions, namely, Col-0 and C24 (Chen, 2013). We propose two possible reasons. One could be that the different molecular bases of heterosis between F_1 hybrids are crossed by different ecotypes. The hybrids of C24 × Col and Col × Ler showed differences in growth vigor at various time points of vegetative development (Groszmann et al., 2014). Another reason could be that circadian rhythm-related genes tend to promote growth through the regulation of transcription. We found that some circadian genes, such as GI and TOC1, are differentially expressed between the parents and hybrids, while a considerable number of the 20 genes involved in the starch biosynthetic process showed only differential m^{6}A methylation rather than differential expression (**Supplementary Table 7**). Since m^{6}A controls RNA fate-related processes, such as mRNA stability, transport, or translation (Dominissini et al., 2012; Meyer et al., 2015; Meyer and Jaffrey, 2017), this study indicates a new layer of regulatory mechanisms contributing to heterosis at the post-transcriptional level in *Arabidopsis*.

Experimental Procedures

Plant Materials and Growth Conditions

Plant materials included two *Arabidopsis* accessions (Col-0, Ler) and their F_1 reciprocal hybrids. F_1 seeds were produced by hand pollination between Col-0 and Ler. Seeds were sown on soil, stratified at 4°C for 3 days to synchronize germination. Plants were then shifted into greenhouse and grown under a long-day condition (16 h in light and 8 h in dark) at 22°C for 21 days. Above-ground tissues were harvested and stored at −80°C for the following experiments.

MeRIP Libraries Construction and Sequencing

MeRIP libraries preparation mainly followed a published procedure (Dominissini et al., 2013). Briefly, total RNA was extracted from leaves in 50 mL conicals using TRIzol (15596018, Ambion, Austin, TX, United States). Poly(A) RNA was enriched (MRN10, Sigma-Aldrich, Saint Louis, MO, United States) and fragmented into ∼100 nt by fragmentation reagent (AM8740, Invitrogen, Carlsbad, CA, United States) for 15 min at 70°C. Few microliters of fragmented RNA was saved as input control, and the left was incubated with m^{6}A antibody (202003, Synaptic Systems, Goettingen, Germany), in 1x IP buffer supplemented with RNasin Plus (N2611, Promega, Madison, WI, United States) for 4 h at 4°C. The antibody-bound RNA was then incubated with pre-blocked protein A beads (10001D, Invitrogen, Carlsbad, CA, United States) at 4°C for 2 h. The immunoprecipitated RNA was released using an elution buffer (1x IP buffer supplemented with 6.7 mM N^6-methyladenosine, M2780, Sigma-Aldrich, Saint Louis, MO, United States). Input and IP libraries were constructed using NEBNext Ultra RNA Library Prep Kit for Illumina (E7645S, NEB, Ipswich, MA, United States) and subjected to sequencing on the Illumina Hiseq X-10 platform.

Reads Pre-processing and Alignment

Raw reads of input and IP samples were processed by trim-galore (version 0.4.1) to remove adaptors and low quality reads and then mapped to the *Arabidopsis* Col-0 reference genome (TAIR 10) using Tophat2 (version 2.1.1) (Kim et al., 2013) with Araport11 annotation in the analyses for parental lines and hybrid lines. We also used Ler reference genome and corresponding annotation (downloaded from NCBI, accession number GCA_001651475.1) to check for possible bias introduced by the reference genome. The parameters were modified (–read-edit-dist 5, –N 5) to obtain more SNP information of Ler and F_1 hybrids. Multiple mapped reads were filtered using the SAMtools package (version 1.9) (Li et al., 2009). Only paired unique reads were used for downstream analysis.

N(6)-Methyladenosine Peak Identification and Annotation

MeTPeak (Cui et al., 2016), a transcriptomic peak caller, was used to identify m^{6}A peaks. In order to get confidence peaks, we maintained peaks on genes with FPKM $\geq$ 1. Moreover, to avoid huge differences in the calculation of peak enrichment due to insufficient coverage, we performed a random sampling of genomic regions and calculated reads of all input samples, and high confidence peaks were selected if the peak region satisfied Input FPKM $\geq$ 5.

To define m^{6}A peak summits, two repeats of input and IP sample were merged, and the coverage of each base of peaks was counted by in-house script (**Supplementary Scripts 1–3**). The residual was calculated by IP reads subtracted by input reads, and the point with the largest residual was referred to as peak summit. The peak summits were intersected with protein-coding gene sequences, which were integrated into a tiered order−3′UTR, 5′ UTR, and CDS, to determine their locations (**Supplementary Script 4**). Additionally, m^{6}A peaks were assigned to start codon and stop codon segments, which was 200 nt centered to start codon and stop codon, respectively, to identify the preference of m^{6}A peaks.

Identification of Differentially Methylated Peaks and Additive/Non-Additive Methylated Peaks

The common m^{6}A peaks between any two samples were defined according to whether they intersected with each other. We calculated read counts of IP and input replicates for each m^{6}A peak of every comparison group (**Supplementary Script 5**). A 2 × 2 contingency table was filled by IP and input normalized reads of samples, respectively. A Fisher's exact test was performed to identify m^{6}A differentially methylated peaks, and *p*-value was adjusted by Bonferroni–Holm correction using R scripts. The differentially methylated peaks should satisfy two requirements: (1) padj <0.05; (2) the difference between any two samples >1.5.

To classify non-additive and additive methylated peaks, Fisher's exact test was performed by comparing the input and IP normalized reads of hybrid and the average of parents' input and IP normalized reads. Only common peaks with padj <0.05 were considered as non-additive methylated peaks. Otherwise, they were referred to as additively methylated peaks.

Identification of Differentially Expressed Genes and Additive/Non-Additive Expressed Genes

The number of reads for each gene was counted using HTSeq (Anders et al., 2015) with a default setting. R package DESeq2 (version 1.22.2) was used for analyzing differentially expressed genes, and only genes with padj < 0.05 were considered as DEGs. If the expression of genes in hybrids was significantly different from mid-parent value (padj < 0.05), these genes were classified as non-additive expressed genes, and the others were referred to as additive expressed genes.

Gene Ontology Analysis

The gene sets were submitted to agriGO database (Tian et al., 2017) to perform GO enrichment analysis. Functional enrichment was performed using the singular enrichment analysis (SEA) tool and TAIR genome locus (TAIR 10) as background. The GO terms with FDR $\leq$ 0.01 were considered to be enriched.

Analysis of Allelic Expression and Allelic N(6)-Methyladenosine Enrichment

To obtain confidence SNPs between Col-0 and Ler, the Ler (downloaded from NCBI) and Col-0 reference genomes (TAIR 10) were cut into 100 bp fragments with 1 bp shift, and then mutually mapped to the reference genome. The read counts of each position were called using the SAMtools "mpileup" command with the parameter "-f." SNPs were first identified if site coverage ≥90X and mutant ratio (mutants/covered reads) ≥90%. The input and IP reads of F_{1CL} and F_{1LC} were separately mapped to the Col-0 reference and the Ler reference, and the reads covered SNPs were calculated. Theoretically, the reads mapped to the corresponding coordinate of the Col-0 and Ler references should be identical, or at least with small bias. Thus, SNPs with severe biased reads (the difference of reads mapped to the corresponding SNPs of two references was more than 10%) were excluded. Additionally, the SNPs that were not homozygous in parent lines were filtered. For allele-specific methylation analysis (**Supplementary Script 6**), we first calculated reads at SNPs within m^{6}A peaks of IP and input replicates of F_1 hybrids, and then filled a 2 × 2 contingency table with normalized reads. A Fisher's exact test was performed to identify allele-specific methylated peaks, and *p*-value was adjusted by Bonferroni-Holm correction using R scripts. Peaks with significant allelic methylation difference (FDR < 0.05) were identified as allele-specific peaks.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. All the sequencing data have been deposited in The National Genomics Data Center (NGDC) under accession number CRA003884.

AUTHOR CONTRIBUTIONS

YW conceived the project. ZX, SD, FZ, and ZW designed the experiments. ZX, XBS, MB, YZ, HW, HX, SW, and HJ

performed the experiments. ZX, XQS, FM, and YW conducted bioinformatics analyses. ZX, SW, and YW wrote the article. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the National Key Research and Development Program of China (2016YFD0101001), Innovative Project of State Key Laboratory for Crop Genetics and Germplasm Enhancement, Fundamental Research Funds for the Central Universities (JCQY201901), Jiangsu Collaborative Innovation Center for Modern Crop Production, and an Innovation and Enterprise Scholar of Jiangsu Province to YW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.685189/full#supplementary-material

Supplementary Figure 1 | Phenotypes of *Arabidopsis* lines and quality of sequencing data. **(A)** The biomass vigor of both F_1 hybrids is higher than that of Col-0 and Ler. Scale bar = 10 mm. Spearman correlations between two biological replicates of input **(B)** mRNA-seq and **(C)** m^{6}A-seq in Col-0, Ler, F_{1CL}, and F_{1LC}.

Supplementary Figure 2 | Global pattern of m^{6}A peaks using the Ler genome sequence as a reference. **(A)** Coverage of normalized reads along transcripts. Each transcript is divided into three non-overlapping features: 5' UTR, CDS, and 3' UTR. **(B)** Distribution of m^{6}A peaks in transcript features of parents and hybrids. **(C)** Relative enrichment of m^{6}A peaks of each transcript feature. Enrichment = Normalized m^{6}A-seq read number divided by normalized input reads of each peak. $^{**}p < 2.2e{-}16$, Wilcoxon rank-sum test.

Supplementary Figure 3 | Features of m^{6}A modifications among the parent lines and F_1 hybrids. **(A)** Number of peaks on transcripts. **(B)** Cumulative plot of m^{6}A methylation enrichment in Col-0, Ler, F_{1CL}, and F_{1LC}. $^{**}p < 2.2e{-}16$, Wilcoxon rank-sum test. **(C)** Number of m^{6}A peaks located at the start codon and stop codon of transcripts. **(D,E)** Number of shared genes containing m^{6}A peaks between accessions.

Supplementary Figure 4 | Diagram of the relationship between m^{6}A methylation level and transcript abundance. **(A)** Number of DEGs between parents and hybrids. **(B)** Number of up- or downregulated DEGs in comparisons of parents/hybrids. **(C)** Scatter plot of DMG-DEGs between accessions showing the relationship of m^{6}A modification and transcript abundance. DMG-DEG indicates DEGs overlapping with DMGs (genes with differentially m^{6}A-methylated peaks). For example, the m^{6}A enrichment ratio of Ler: F_{1CL} is calculated as log2 (enrichment of Ler/enrichment of F_{1CL}) of m^{6}A peaks. The gene expression ratio of Ler: F_{1CL} is calculated as log2 (FPKM of Ler/FPKM of F_{1CL}) of transcripts. *n* indicates the number of DMG-DEGs in each quadrant. Gene m^{6}A enrichment is calculated by normalized m^{6}A-seq reads number divided by normalized input reads of peaks within the transcript, and gene expression is indicated by the FPKM of the input RNA-seq data.

Supplementary Figure 5 | Enriched biological functions of differentially m^{6}A-methylated genes. **(A)** Enriched GO terms of genes associated with differentially m^{6}A-methylated peaks (DMPs) that are not DEGs in comparisons of parents/hybrids. **(B)** Enriched GO terms of DEGs not associated with differentially m^{6}A-methylated peaks (DMPs) in comparisons of parents/hybrids. Only some of the enriched GO terms enriched in the comparisons between parents/hybrids rather than Col/Ler are shown in the figure. All the GO terms are listed in **Supplementary Table 4**.

Supplementary Table 1 | List of m^{6}A peaks identified in all the samples.

Supplementary Table 2 | List of DMPs between the samples.

Supplementary Table 3 | List of non-additive and additive expressed genes and m^{6}A peaks in hybrids.

Supplementary Table 4 | The results of GO analysis of DMPs in each comparison.

Supplementary Table 5 | Shared GO terms of any two samples.

Supplementary Table 6 | Genes related to growth vigor.

Supplementary Table 7 | m^{6}A methylation of starch biosynthetic process and circadian rhythm-related genes.

REFERENCES

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Anderson, S. J., Kramer, M. C., Gosai, S. J., Yu, X., Vandivier, L. E., Nelson, A. D. L., et al. (2018). N6-methyladenosine inhibits local ribonucleolytic cleavage to stabilize mRNAs in *Arabidopsis*. *Cell Rep*. 25, 1146–1157. doi: 10.1016/j.celrep.2018.10.020

Arribas-Hernandez, L., Bressendorff, S., Hansen, M. H., Poulsen, C., Erdmann, S., and Brodersen, P. (2018). An m6A-YTH module controls developmental timing and morphogenesis in *Arabidopsis*. *Plant Cell* 30, 952–967. doi: 10.1105/tpc.17.00833

Birchler, J. A. (2015). Heterosis: the genetic basis of hybrid vigour. *Nat. Plants* 1, 15020. doi: 10.1038/nplants.2015.20

Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003). In search of the molecular basis of heterosis. *Plant Cell* 15, 2236–2239. doi: 10.1105/tpc.151030

Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., and Veitia, R. A. (2010). Heterosis. *Plant Cell* 22, 2105–2112. doi: 10.1105/tpc.110.076133

Bodi, Z., Zhong, S., Mehra, S., Song, J., Graham, N., Li, H., et al. (2012). Adenosine methylation in *Arabidopsis* mRNA is associated with the 3' end and reduced levels cause developmental defects. *Front. Plant Sci.* 3:48. doi: 10.3389/fpls.2012.00048

Chen, Z. J. (2010). Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* 15, 57–71. doi: 10.1016/j.tplants.2009.12.003

Chen, Z. J. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet*. 14, 471–482. doi: 10.1038/nrg3503

Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., et al. (2018). mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* 561, 556–560. doi: 10.1038/s41586-018-0538-8

Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401, 157–161. doi: 10.1038/43657

Cui, X., Meng, J., Zhang, S., Chen, Y., and Huang, Y. (2016). A novel algorithm for calling mRNA m6A peaks by modeling biological variances in MeRIP-seq data. *Bioinformatics* 32, i378–i385. doi: 10.1093/bioinformatics/btw281

Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N6-methyladenosine by m6A-seq based on immunocapturing and massively parallel sequencing. *Nat. Protoc.* 8, 176–189. doi: 10.1038/nprot.2012.148

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485, 201–206. doi: 10.1038/nature11112

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M. S., et al. (2016). The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 530, 441–446. doi: 10.1038/nature16998

Dong, X., Reimer, J., Göbel, U., Engelhorn, J., He, F., Schoof, H., et al. (2012). Natural variation of H3K27me3 distribution between two *Arabidopsis* accessions and its association with flanking transposable elements. *Genome Biol.* 13:R117. doi: 10.1186/gb-2012-13-12-r117

Duan, H.-C., Wei, L.-H., Zhang, C., Wang, Y., Chen, L., Lu, Z., et al. (2017). ALKBH10B is an RNA N6-methyladenosine demethylase affecting *Arabidopsis* floral transition. *Plant Cell* 29, 2995–3011. doi: 10.1105/tpc.16.00912

Gilbert, W. V., Bell, T. A., and Schaening, C. (2016). Messenger RNA modifications: form, distribution, and function. *Science* 352:1408. doi: 10.1126/science.aad8711

Greaves, I. K., Groszmann, M., Wang, A., Peacock, W. J., and Dennis, E. S. (2014). Inheritance of trans chromosomal methylation patterns from *Arabidopsis* F1 hybrids. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2017–2022. doi: 10.1073/pnas.1323656111

Groszmann, M., Gonzalez-Bayon, R., Greaves, I. K., Wang, L., Huen, A. K., Peacock, W. J., et al. (2014). Intraspecific *Arabidopsis* hybrids show different patterns of heterosis despite the close relatedness of the parental genomes. *Plant Physiol*. 166, 265–280. doi: 10.1104/pp.114.243998

He, G., Chen, B., Wang, X., Li, X., Li, J., He, H., et al. (2013). Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. *Genome Biol*. 14:R57. doi: 10.1186/gb-2013-14-6-r57

He, G., Zhu, X., Elling, A. A., Chen, L., Wang, X., Guo, L., et al. (2010). Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22, 17–33. doi: 10.1105/tpc.109.072041

Hochholdinger, F., and Hoecker, N. (2007). Towards the molecular basis of heterosis. *Trends Plant Sci*. 12, 427–432. doi: 10.1016/j.tplants.2007.08.005

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol*. 7:885. doi: 10.1038/nchembio.687

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). Tophat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 14:R36. doi: 10.1186/gb-2013-14-4-r36

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, Q., Li, X., Tang, H., Jiang, B., Dou, Y., Gorospe, M., et al. (2017). NSUN2-mediated m5C methylation and METTL3/METTL14-mediated m6A methylation cooperatively enhance p21 translation. *J. Cell. Biochem*. 118, 2587–2598. doi: 10.1002/jcb.25957

Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., et al. (2016). Transcriptome-wide mapping reveals reversible and dynamic N1-methyladenosine methylome. *Nat. Chem. Biol*. 12, 311–316. doi: 10.1038/nchembio.2040

Li, Y., Wang, X., Li, C., Hu, S., Yu, J., and Song, S. (2014). Transcriptome-wide N6-methyladenosine profiling of rice callus and leaf reveals the presence of tissue-specific competitors involved in selective mRNA modification. *RNA Biol*. 11, 1180–1188. doi: 10.4161/rna.36281

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N6-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564. doi: 10.1038/nature14234

Luo, G.-Z., MacQueen, A., Zheng, G., Duan, H., Dore, L. C., Lu, Z., et al. (2014). Unique features of the m6A methylome in *Arabidopsis thaliana*. *Nat. Commun*. 5, 5630–5630. doi: 10.1038/ncomms6630

Luo, J., Wang, Y., Wang, M., Zhang, L., Peng, H., Zhou, Y., et al. (2020). Natural variation in RNA m6A methylatioin and its relationship with translational status. *Plant Physiol*. 182, 332–344. doi: 10.1104/pp.19.00987

Luo, J. H., Wang, M., Jia, G. F., and He, Y. (2021). Transcriptome-wide analysis of epitranscriptome and translational efficiency associated with heterosis in maize. *J. Exp. Bot*. 72, 2933–2946. doi: 10.1093/jxb/erab074

Lv, Z., Zhang, W., Wu, Y., Huang, S., Zhou, Y., Zhang, A., et al. (2019). Extensive allele-level remodeling of histone methylation modification in reciprocal F1 hybrids of rice subspecies. *Plant J*. 97, 571–586. doi: 10.1111/tpj.14143

Manning, K., Tor, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., et al. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet*. 38, 948–952. doi: 10.1038/ng1841

Meyer, K. D., and Jaffrey, S. R. (2017). Rethinking m6A readers, writers, and erasers. *Annu. Rev. Cell Dev. Biol*. 33, 319–342. doi: 10.1146/annurev-cellbio-100616-060758

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., et al. (2015). 5' UTR m6A promotes cap-independent translation. *Cell* 163, 999–1010. doi: 10.1016/j.cell.2015.10.012

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149, 1635–1646. doi: 10.1016/j.cell.2012.05.003

Miao, Z., Zhang, T., Qi, Y., Song, J., Han, Z., and Ma, C. (2020). Evolution of the RNA N6-methyladenosine methylome mediated by genomic duplication. *Plant Physiol*. 182, 345–360. doi: 10.1104/pp.19.00323

Moghaddam, A. M., Roudier, F., Seifert, M., Berard, C., Magniette, M. L., Ashtiyani, R. K., et al. (2011). Additive inheritance of histone modifications in *Arabidopsis thaliana* intra-specific hybrids. *Plant J*. 67, 691–700. doi: 10.1111/j.1365-313X.2011.04628.x

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

Scutenaire, J., Deragon, J. M., Jean, V., Benhamed, M., Raynaud, C., Favory, J. J., et al. (2018). The YTH domain protein ECT2 is an m6A reader required for normal trichome branching in *Arabidopsis*. *Plant Cell* 30, 986–1005. doi: 10.1105/tpc.17.00854

Shen, H., He, H., Li, J., Chen, W., Wang, X., Guo, L., et al. (2012). Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* 24, 875–892. doi: 10.1105/tpc.111.094870

Shen, L., Liang, Z., Gu, X., Chen, Y., Teo, Z. W., Hou, X., et al. (2016). N6-methyladenosine RNA modification regulates shoot stem cell fate in *Arabidopsis*. *Dev. Cell* 38, 186–200. doi: 10.1016/j.devcel.2016.06.008

Shindo, C., Lister, C., Crevillen, P., Nordborg, M., and Dean, C. (2006). Variation in the epigenetic silencing of FLC contributes to natural variation in *Arabidopsis* vernalization response. *Genes Dev*. 20, 3079–3083. doi: 10.1101/gad.405306

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J*. 37, 914–939. doi: 10.1111/j.1365-313X.2004.02016.x

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res*. 45, W122–W129. doi: 10.1093/nar/gkx382

Wan, Y., Tang, K., Zhang, D., Xie, S., Zhu, X., Wang, Z., et al. (2015). Transcriptome-wide high-throughput deep m6A-seq reveals unique differential m6A methylation patterns between three organs in *Arabidopsis thaliana*. *Genome Biol*. 16:272. doi: 10.1186/s13059-015-0839-2

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* 505, 117–120. doi: 10.1038/nature12730

Wei, L.-H., Song, P., Wang, Y., Lu, Z., Tang, Q., Yu, Q., et al. (2018). The m6A reader ECT2 controls trichome morphology by affecting mRNA stability in *Arabidopsis*. *Plant Cell* 30, 968–985. doi: 10.1105/tpc.17.00934

Yang, M., Wang, X., Huang, H., Ren, D., Su, Y., Zhu, P., et al. (2016). Natural variation of H3K27me3 modification in two *Arabidopsis* accessions and their hybrid. *J. Integr. Plant Biol*. 58, 466–474. doi: 10.1111/jipb.12443

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., et al. (2008). MTA is an *Arabidopsis* messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. *Plant Cell* 20, 1278–1288. doi: 10.1105/tpc.108.058883

Zhou, L., Tian, S., and Qin, G. (2019). RNA methylomes reveal the m6A-mediated regulation of DNA demethylase gene *SlDML2* in tomato fruit ripening. *Genome Biol*. 20:156. doi: 10.1186/s13059-019-1771-7

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

frontiers in Plant Science

MINI REVIEW
published: 17 June 2021
doi: 10.3389/fpls.2021.687416

Removal of H3K27me3 by JMJ Proteins Controls Plant Development and Environmental Responses in *Arabidopsis*

*Nobutoshi Yamaguchi**

Division of Biological Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

OPEN ACCESS

Edited by:
Stewart Gillmor, National Laboratory of Genomics for Biodiversity, Center for Research and Advanced Studies, National Polytechnic Institute of Mexico (CINVESTAV), Mexico

Reviewed by:
Gerardo del Toro, Swedish University of Agricultural Sciences, Sweden
Hongchang Cui, Florida State University, United States

***Correspondence:**
Nobutoshi Yamaguchi
nobuy@bs.naist.jp

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

Received: *29 March 2021*
Accepted: *26 May 2021*
Published: *17 June 2021*

Citation:
Yamaguchi N (2021) Removal of H3K27me3 by JMJ Proteins Controls Plant Development and Environmental Responses in Arabidopsis. Front. Plant Sci. 12:687416. doi: 10.3389/fpls.2021.687416

Trimethylation of histone H3 lysine 27 (H3K27me3) is a highly conserved repressive histone modification that signifies transcriptional repression in plants and animals. In *Arabidopsis thaliana*, the demethylation of H3K27 is regulated by a group of *JUMONJI DOMAIN-CONTANING PROTEIN* (*JMJ*) genes. Transcription of *JMJ* genes is spatiotemporally regulated during plant development and in response to the environment. Once *JMJ* genes are transcribed, recruitment of JMJs to target genes, followed by demethylation of H3K27, is critically important for the precise control of gene expression. JMJs function synergistically and antagonistically with transcription factors and/or other epigenetic regulators on chromatin. This review summarizes the latest advances in our understanding of *Arabidopsis* H3K27me3 demethylases that provide robust and flexible epigenetic regulation of gene expression to direct appropriate development and environmental responses in plants.

Keywords: *Arabidopsis*, development, demethylases, epigenetics, environmental response, JUMONJI, histone modification, H3K27me3

INTRODUCTION

Chromatin is critically important for gene expression during plant development and in response to the environment (Eccleston et al., 2013; Bruneau et al., 2019). Chromatin is composed of genomic DNA, histones, and accessory proteins, with approximately 150 base pairs of DNA wrapped around each octameric histone protein complex (Vergara and Gutierrez, 2017; van Steensel and Furlong, 2019). Each histone protein consists of a structural core at the C terminus and an unstructured tail domain at the N terminus. The N-terminal flexible histone tails often possess extensive posttranslational modifications, such as acetylation, methylation, and ubiquitination on lysine residues, methylation and citrullination on arginine residues, and phosphorylation of serine, threonine, and tyrosine residues. These modifications cause epigenetic changes in chromatin and lead to changes in gene expression.

One chromatin modification, trimethylation of histone H3 lysine 27 (H3K27me3), mediates epigenetic silencing of gene expression (Xiao and Wagner, 2015; Xiao et al., 2016). In general, H3K27me3 marks occur within facultative heterochromatin, in which gene expression is repressed but can be activated in response to developmental or environmental cues. In animals and plants, H3K27me3 deposition and removal are mediated by specific enzymes termed "writers" and "erasers", respectively. Polycomb repressive complex 2 (PRC2), a multisubunit epigenetic repressor complex, writes H3K27me3 marks associated with gene repression. By contrast, histone demethylases, such as the Jumonji C (JmjC)-containing eraser demethylases, can demethylate H3K27me3 and thereby counteract the action of writer methylases (Crevillén, 2020). Understanding the role of JmjC-containing demethylases is crucial to understanding the effects of H3K27me3 in plant development and environmental responses. Although PRC2 and its actions have been reasonably well characterized through decades of research, existing knowledge about H3K27me3 demethylases and demethylation is relatively limited. In the last decade, however, research on H3K27me3 removal in *Arabidopsis thaliana* (*Arabidopsis*) has made great progress. To date, five JMJ proteins have been identified as H3K27me3 demethylases: EARLY FLOWERING 6 (ELF6)/JUMONJI DOMAIN-CONTAINING PROTEIN11 (JMJ11), RELATIVE OF ELF6 (REF6)/JMJ12, JMJ13, JMJ30, and JMJ32 (Lu F. et al., 2011; Crevillén et al., 2014; Gan et al., 2014; Cui et al., 2016; Yan et al., 2018). Here, we summarize current understanding of a group of JmjC-containing demethylases of H3K27me3, with emphasis on the most recent advances in knowledge.

H3K27 DEMETHYLASES GOVERN MANY PROCESSES IN PLANT LIFE

Upon sensing developmental or environmental cues, JMJ proteins make genomic regions accessible by removing repressive H3K27me3 marks to generate a legible genome that is specific to a particular cell type, developmental stage, or environmental condition. Functional analysis of loss-of-function *jmj* mutants in Arabidopsis has indicated that JMJ proteins make major contributions to developmentally or environmentally triggered transcriptional reprogramming events. REF6, ELF6, and JMJ13 make a broader contribution to plant growth and development than JMJ30 and JMJ32, which play more specific and redundant roles in environmental responses.

H3K27 Demethylases Accumulate in Various Tissues

The divergence in the biological roles of H3K27me3 demethylases might be due to their different spatial and temporal expression patterns. The spatial distribution of REF6, ELF6, JMJ13, JMJ30, and JMJ32 proteins was examined by introducing constructs harboring their upstream and coding sequences fused with sequences encoding the β-glucuronidase (GUS) reporter into wild-type plants (Noh et al., 2004; Gan et al., 2014; Zheng et al., 2019). Among these five GUS reporters, JMJ30-GUS highly accumulated in various plant organs, such as leaves, roots, and flowers (Gan et al., 2014). REF6-GUS, JMJ13-GUS, and JMJ32-GUS show moderate accumulation in young leaves near the shoot apical meristem and in root tips but lower accumulation in the leaf vasculature (Noh et al., 2004; Gan et al., 2014; Zheng et al., 2019). By contrast, ELF6-GUS accumulates only in the distal part of young leaves. Current spatial expression data were obtained mainly by whole-mount GUS staining. Our understanding of JMJ accumulation is still limited largely to the organ level. Expression analysis derived from GUS staining may not be precise, due to diffusion of the enzyme outside the tissue, as compared with fluorescent protein–based experiments.

Transcriptome data from publicly available databases increase the understanding of demethylase function in Arabidopsis and allow functions to be inferred. Shoot apex–specific RNA sequencing (RNA-seq) and cell type–specific single-cell (sc) RNA-seq data revealed different expression patterns for the six *Arabidopsis* H3K27me3 demethylase genes (Winter et al., 2007; Ryu et al., 2019; Tian et al., 2019). In the shoot apical meristem, *JMJ30* is highly expressed, whereas the other genes are weakly expressed (Tian et al., 2019). Among eight different domains within the shoot apical meristem, *JMJ30* expression is higher in the *CLAVATA3* (*CLV3*) and *ARABIDOPSIS THALIANA MERISTEM LAYER 1* (*ATML1*) expression domains (Lu et al., 1996; Brand et al., 2002). Because *CLV3* and *ATML1* are specifically expressed in the central zone and layer 1, respectively, these observations suggest that *JMJ30* is also highly expressed in the center and/or epidermis of the shoot apical meristem. In the root, *JMJ13*, *ELF6*, *JMJ30*, and *RFF6* are expressed in a cell type–specific manner (Ryu et al., 2019). High *JMJ13* expression in the protoxylem suggests that it has a specific function in this tissue. These high-resolution differential expression patterns suggest that histone demethylation is tissue or cell type specific. Expression specificity at the cell-type or cellular levels needs to be characterized in detail to further our understanding of when and where JMJ proteins work.

Although epigenetic regulation is thought to be important in responses to environmental stimuli, few reports have described the relationship between environmental stress and the induction of *JMJ* genes. *JMJ30* expression is further enhanced by the stress hormone abscisic acid (ABA) and by salt stress, drought stress, and heat stress compared to control conditions (Qian et al., 2015; Wu et al., 2019a; Yamaguchi et al., 2020). ABA treatment triggers a rapid increase in JMJ30 protein levels but does not change the area of JMJ30 expression, based on whole-mount GUS staining (Wu et al., 2019a). The expression of *JMJ13* is affected by light and temperature conditions, according to GUS expression data (Zheng et al., 2019). *REF6* expression is induced by long-term heat exposure (Liu et al., 2019). To date, no effects of environmental stress on the regulation of *ELF6* and *JMJ32* expression have been reported. Bulk transcriptome datasets also largely support these results (Qian et al., 2015). Expression specificity and subcellular localization of JMJ proteins in response to environmental stimuli should also be addressed with higher resolution in the future.

Mutant Phenotypes and Key Targets of H3K27 Demethylases

The seeds of the *ref6* mutant germinate later than wild type (Li et al., 2016; Chen et al., 2020). REF6 induces two key genes for ABA catabolism, *CYP707A1* and *CYP707A3*, through removal of H3K27me3. *CYP707A1* and *CYP707A2* encode ABA 8′-hydeoxylases and play key roles in reducing ABA levels (Okamoto et al., 2006). Overexpression of *CYP707A1* by cauliflower mosaic virus (CaMV) 35S promoter rescues the dormancy phenotype of the *ref6* mutant (Chen et al., 2020). The *jmj30 jmj32* double mutant, by contrast, shows no difference in seed dormancy phenotype from wild type (Wu et al., 2019a).

Under normal growth conditions, *ref6* and *elf6* mutants have similar leaf phenotypes that include reduced petiole length, which is characteristic of brassinosteroid (BR)-defective mutants (Yu et al., 2008). A shorter leaf blade is seen in *ref6* but not *elf6* plants, suggesting that the REF6 and ELF6 proteins have tissue-specific roles. The *ref6* mutation further enhances the phenotype of a BR-deficient mutant. In the *ref6 elf6* double mutant, expression of BR-regulated genes, such as *TOUCH 4* (*TCH4*), is reduced. Later in leaf development, *ref6* delays chlorophyll degradation (Wang et al., 2019) and REF6 promotes general leaf senescence by directly activating senescence-related genes, including *ETHYLENE INSENSITIVE 2* (*EIN2*), *OLEOSIN 1* (*ORE1*), and *NONYELLOWING* genes (*NYE*s). The *jmj13* single mutant does not display detectable abnormalities in leaf phenotype (Zheng et al., 2019), but the *ref6 elf6 jmj13* triple mutant has shorter petioles than *ref6 elf6*, suggesting that REF6, ELF6, and JMJ13 are essential developmental regulators (Yan et al., 2018). The *jmj30 jmj32* double mutant, by contrast, shows no difference in leaf phenotype from wild type (Yamaguchi et al., 2020).

All five Arabidopsis H3K27me3 demethylases regulate flowering time, but in distinct fashions. *REF6* and *ELF6* were originally identified on the basis of their influence on flowering-time phenotypes: under long-day conditions, *ref6* mutants are late flowering and *elf6* and *jmj13* mutants are early flowering (Noh et al., 2004; Zheng et al., 2019). REF6 directly induces floral activator genes, such as *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*) and *FRUITFULL* (*FUL*) (Hou et al., 2014; Hyun et al., 2016). ELF6 binds to the regulatory region of the floral repressor gene *FLOWERING LOCUS C* (*FLC*) for transcriptional activation (Yang et al., 2016). In *jmj13* mutants, the floral repressor gene *SHORT VEGETATIVE PHASE* (*SVP*) is downregulated. By contrast, flowering-time defects in *jmj30 jmj32* are observed only at high ambient temperatures but not under long-day conditions (Gan et al., 2014; Yan et al., 2014). Thus, the H3K27me3 demethylases *REF6*, *ELF6*, *JMJ13*, *JMJ30*, and *JMJ32* show differences as well as similarities in how they influence flowering time.

Differences are also observed between the phenotypes of *elf6* and *jmj13* during flower development (Keyzor et al., 2021). In wild type and the *ref6* mutant, the initial one or two flowers do not undergo self-pollination and form very short fruits without seeds; *elf6* plants display increased self-fertility and consistent fruit production. Conversely, the *jmj13* mutant shows reduced fertility and gives rise to aborted fruits up to the eighth flower on the primary inflorescence. *JASMONATE-ZIM-DOMAIN PROTEIN 7* (*JAZ7*), *SMALL AUXIN UP RNA 26* (*SAUR26*) and *ARABINOGALACTAN PROTEINs* (*AGPs*) are downregulated in *jmj13* buds. No defects in floral developmental have been reported for *jmj30* and *jmj32* mutants.

The functions of JMJ30 and JMJ32 appear to be relatively distinct from those of REF6, ELF6, and JMJ13. *jmj30* mutants show a circadian phenotype (Jones et al., 2010), and the *JMJ30* gene was originally identified due to its co-expression with *TIMING OF CAB1 EXPRESSION 1* (*TOC1*). Consistent with the circadian oscillation in *JMJ30* expression, circadian rhythms in reporter-gene activity in *jmj30* mutants are significantly shorter than those in wild type. *JMJ30* and *TOC1* interact genetically to promote the expression of *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*). By contrast, no circadian oscillation in *JMJ32* expression is observed (Lu S. X. et al., 2011), suggesting that *JMJ30* and *JMJ32* are regulated by distinct mechanisms.

jmj30 mutants also feature phenotypes that are dependent on environmental conditions. Callus formation induced by incubating leaf explants on callus-inducing medium is reduced in *jmj30* mutants. JMJ30 promotes the expression of *LATERAL ORGAN BOUNDARIES DOMAIN 16* (*LBD16*) and *LBD29* to establish root primordium–like unorganized cell masses (Lee et al., 2018). Furthermore, stress hormone–induced growth arrest is compromised in *jmj30 jmj32* double mutants (Wu et al., 2019a,b, 2020), and JMJ30 directly activates *SNF1-RELATED PROTEIN KINASE 2.8* (*SnRK2.8*) and *BRASSINAZOLE RESISTANT1* (*BZR1*) to maintain a balance between stress responses and growth. Acquired thermotolerance is also reduced in *jmj30 jmj32 ref6 elf6* quadruple mutants (Yamaguchi et al., 2020). JMJ30 binds to *HEAT SHOCK PROTEIN 17.6C* (*HSP17.6C*) and *HSP22* and activates their transcription in response to heat.

Although the interactions between JMJ proteins and downstream targets is regulated in a spatiotemporal manner, it is not yet known how exactly JMJ proteins lead to H3K27me3 removal. Most phenotyping has been conducted in knock-out or knock-down mutants, while mutant rescue by expressing downstream targets has used CaMV35S-based overexpression lines (Wu et al., 2019a; Chen et al., 2020; Yamaguchi et al., 2020). To assess the precise roles of JMJ during plant development and environmental responses, conditional *jmj* mutants should be employed. Furthermore, organ-, tissue-, or cell type–specific phenotypic rescues using appropriate promoters are required to understand when and where JMJ proteins function.

PROTEIN STRUCTURE AND CHROMATIN-TARGETING MECHANISMS OF H3K27 DEMETHYLASES

Phylogenetic analysis of JmjC-containing demethylases defined 14 subfamilies and identified more than 10 members in land plants. Green algae such as *Chlamydomonas* and *Volvox* include only two members of this family, implying that the functions of

JMJ proteins may have been important for plant adaptation to land (Qian et al., 2015). Arabidopsis contains 21 JMJ proteins (Lu et al., 2008). Although not all family members have been fully characterized, they include putative H3K9me3-, H3K36me3-, H3K4me3-, and H3K27me3-specific demethylases. Additional H3K27me3 demethylases may also exist. ELF6 and REF6 show highest sequence similarities to the H3K9me3- and H3K36me3-specific KMD4 demethylases. The precise functions of the remaining JMJ proteins need to be carefully examined in a manner that is unbiased by sequence similarity.

The ELF6, REF6, and JMJ13 proteins belong to the plant-specific KMD4 subfamily, which is present in land plants but not in green algae (Lu et al., 2008; Qian et al., 2015). REF6 contains JmjN, JmjC, and C2H2-type zinc-finger (ZnF) domains (**Figure 1A**). The REF6 protein characteristically possesses four tandem repeats of the ZnF domain. These domains are essential for REF6 function, as complementation of the *ref6* mutant through the introduction of REF6 without a ZnF domain fails to rescue the *ref6* mutant phenotype. Some histone demethylases, such as KDM2, interact with chromatin via direct binding to DNA through ZnF domains. ZnF is one of the largest class of DNA-binding domains. Consistent with this, REF6 functions as a DNA sequence–specific H3K27me3 demethylase. Genome-wide REF6 binding studies and crystal structure analysis revealed that the ZnF domains of REF6 recognize the CTCTGYTY DNA motif for H3K27me3 removal (**Figures 1B**, **2A**) (Cui et al., 2016; Li et al., 2016; Tian et al., 2020). The ZnF domains of REF6 complex with *NAC004* double-stranded (ds) DNA by forming a half-cross-braced structure (**Figure 1B**). Interactions at the interface between REF6 and dsDNA, such as hydrogen bonds, electrostatic interactions, and hydrophobic interactions, strengthen their binding. dsDNA binding induces profound conformation changes (Tian et al., 2020). Conformational plasticity of DNA allows REF6 to recognize diverse target genes.

EARLY FLOWERING 6 is the closest homolog of REF6. Those two proteins share a high sequence similarity. How ELF6

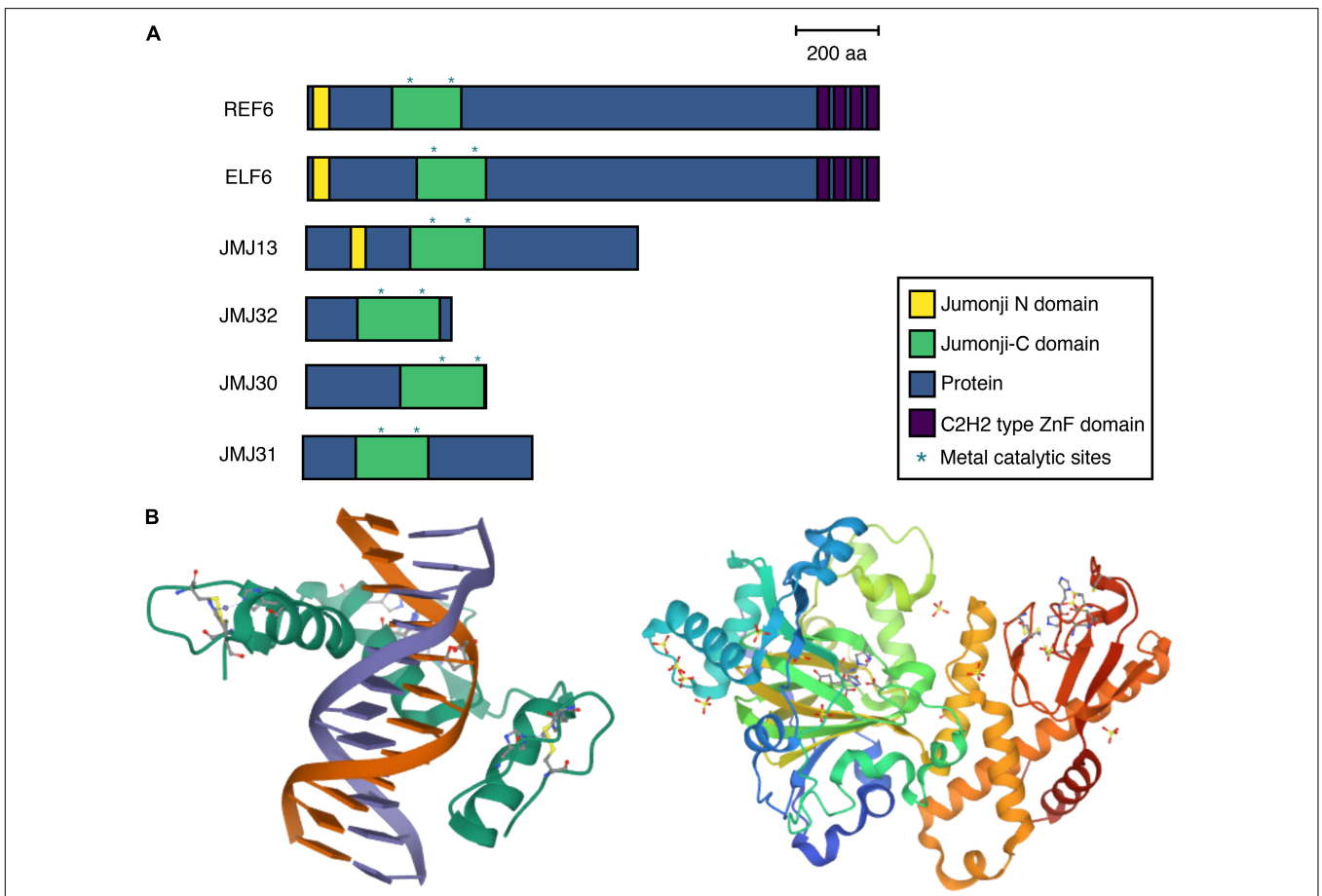

FIGURE 1 | Conserved motifs, and three-dimensional structure of H3K27me3 demethylases in *Arabidopsis thaliana*. **(A)** The domain structures of Jumonji domain–containing proteins (JMJs) in *A. thaliana*. The positions of the Jumonji N domain, Jumonji C domain, protein, and C2H2-type zinc-finger (ZnF) domains are indicated in yellow, green, blue, and purple, respectively; metal catalytic sites are indicated with asterisks. Scale bar = 200 amino acids. **(B)** Crystal structure of JMJ proteins. *Left* REF6 ZnF and *NAC004*-mC3 double-stranded (ds) DNA. Ribbon representation of REF6-DNA structure. REF6 protein is shown in green, while dsDNA is shown in orange and purple. α-helices and β-sheets are represented by spiral ribbons and green arrows, respectively. A few residues engage Zn^{2+} ion. *Right* The JMJ13 catalytic domain in complex with AKG. Ribbon representation of JMJ13 structure. Blue and green ribbons represent JMJ domains. Orange and red ribbons show helical and ZnF domains, respectively. The data were obtained from the Protein Data Bank (https://biorender.com).

recognizes DNA is currently unknown. Because the mutant phenotypes of *elf6* and *ref6* differ, the recognition motifs or mechanisms might also differ between REF6 and ELF; a recent study showed that REF6 and ELF6 play distinct roles in H3K27me3 and H3K27me1 homeostasis (Antunez-Sanchez et al., 2020). ELF6 regulates a small subset of genes compared to REF6. This could be due to less protein structural plasticity of ELF6 and/or a difference in DNA-binding affinity. Further analysis, such as determination of crystal structures of ELF6-DNA complexes, is required to reveal the structural basis for the epigenetic modification recognition.

Although JMJ13, like REF6, belongs to the KMD4 subfamily and functions to remove H3K27me3 *in vitro* and *in vivo*, it possesses a different DNA-recognition mechanism. JMJ13 does not contain ZnF domains at the C terminus (**Figure 1A**); instead, its catalytic domain (JMJ13CD) contains Jmj and helical domains, as well as a unique C4HCHC-type ZnF domain. Crystal structure analysis using JMJ13CD revealed that JMJ13 recognizes the H3K27me3 peptide and functions as a reader of the histone modification state (**Figures 1B**, **2B**) (Zheng et al., 2019). The interactions between JMJ13 and the H3K27me3 peptide are restricted to the region between H3R26 and H3P30. Because other JMJ proteins are predicted to possess different putative ZnF domains, such as the C5HC2-type, the recognition of histone modifications by ZnF domains located within the JmjC domain might also occur for other JMJ proteins.

JMJ30 and JMJ32, as well as their close homolog JMJ31, belong to the JmjC-domain-only group. Consistent with their domain structure, two clades of protein homology are identified by phylogenetic analysis: one contains ELF6, REF6, and JMJ13 and the other contains JMJ30, JMJ31, and JMJ32 (Lu et al., 2008; Qian et al., 2015). Although the function and regulation of JMJ30 and JMJ32 are relatively well characterized, nothing is known about their three-dimensional protein structure or DNA–histone recognition mechanisms. As is often observed for REF6 and ELF6 recruitment (**Figure 2C**), JMJ30 physically interacts with tissue-specific transcription factors, such as EARLY FLOWERING MYB PROTEIN (EFM) and AUXIN RESPONSE FACTORs (ARFs) (Yan et al., 2014; Lee et al., 2018). Furthermore, JMJ30 activity affects H3K9me3 and H3K36me3 in addition to H3K27me3. Further experiments, such as JMJ30 chromatin immunoprecipitation and deep sequencing (ChIP-seq), are required to precisely understand their biochemical functions.

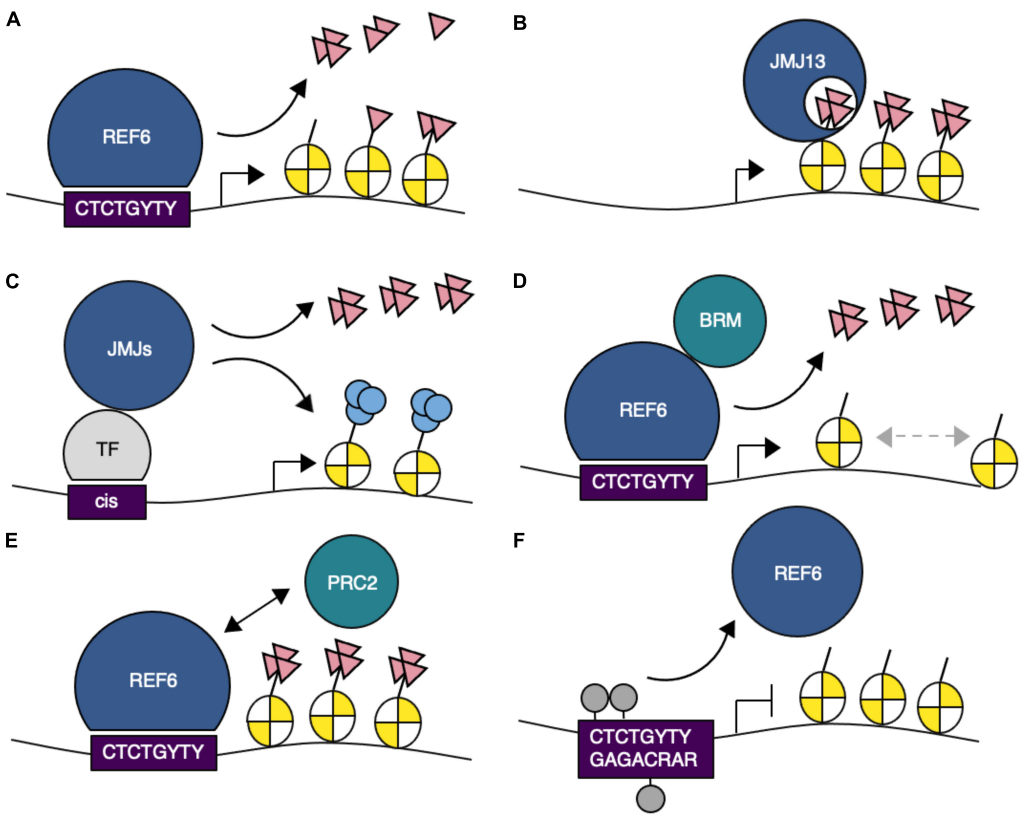

FIGURE 2 | Schematic representation of the function of H3K27me3 demethylases in gene regulation. **(A)** REF6 demethylates H3K27me1/2/3 via recognition of CTCTGYTY (where Y is C or T) DNA motifs. **(B)** JMJ13 recognizes H3K27me3 marks. **(C)** JMJ proteins interact with transcription factors and are recruited to their target sites. **(D)** REF6 interacts with BRAHMA (BRM) and mediates nucleosome positioning. **(E)** REF6 prevents the uncontrolled spreading of PRC2-mediated chromatin silencing. **(F)** DNA methylation at CTCTGYTY motifs prevents REF6 targeting. H3K27me3, pink triangles; H3K9ac, light blue circles; DNA methylation, gray circles.

Genome-wide JMJ protein binding and histone modification data were obtained by ChIP-seq. Since ChIP-seq assays require large numbers of input cells/tissues, whole plants are often used for the assays. Hence, spatial information is completely lost. Recently, low-input binding tests in plants, such as CUT&Tag, CUT&RUN, nCUT&Tag, and ChIL, have been developed to study interactions between DNA and proteins using low-input samples or single live cells (Zheng and Gehring, 2019; Sakamoto et al., 2020; Tao et al., 2020; Ouyang et al., 2021). By combining these with cell sorting or laser microdissection techniques, cell type–specific JMJ protein binding and histone modification data can be obtained in the future. These analyses may contribute to our understanding of the precise spatiotemporal regulation of H3K27me3 demethylation.

INTERACTIONS BETWEEN H3K27 DEMETHYLASES AND OTHER FACTORS ON CHROMATIN

Genome-wide binding analysis coupled with immunoprecipitation and mass spectrometry (IP-MS) identified an interaction between REF6 and the SWI/SNF-type chromatin remodeling ATPase BRAHMA (BRM) on chromatin *in vivo* (Li et al., 2016). REF6 and BRM bind to many common genomic loci that contain CTCTGYTY motifs. Recruitment of BRM to target loci is dependent on REF6 function, but REF6 does not require BRM activity for its own targeting. Thus, REF6 directly binds to chromatin containing the CTCTGYTY motifs and subsequently recruits BRM to activate targets, potentially through changes in nucleosome position (**Figure 2D**).

An antagonistic role between REF6 or BRM and PRC2 at their target loci has been demonstrated (Bezhani et al., 2007; Lu F. et al., 2011; Wu et al., 2012; Li et al., 2015). Antagonism is often mediated by competitive binding at the same sites on chromatin (Zhu et al., 2020). However, PRC2 preferentially binds to different motifs, such as the telobox and GAGA motifs (Hecker et al., 2015; Xiao et al., 2017; Zhou et al., 2018), suggesting that competitive antagonism is unlikely to occur between REF6/BRM and PRC2; moreover, the binding patterns of PRC2 and REF6 do not overlap. REF6 is localized to the boundaries of H3K27me3 regions, which are covered by PRC2 (Yan et al., 2018) (**Figure 2E**). The spreading of H3K27me3 observed in *ref6 elf6 jmj13* triple mutants indicates that the function of REF6 binding inhibits the spreading of H3K27me3, but how ELF6 and JMJ13 contribute to preventing this spreading remains unclear.

Recognition of dsDNA by REF6 not only relies on DNA sequence but also is affected by DNA methylation and sequence-dependent conformations of DNA (Qiu et al., 2019). REF6 preferentially binds to hypomethylated CTCTGYTY motifs. Methylation of CHG within the motif attenuates REF6-binding affinity (**Figure 2F**), and the minor groove width of each nucleotide in the structure of the complex differs considerably. This difference affects recognition of the CTCTGYTY motifs by REF6 and its binding affinity in *CUP-SHAPED COTYLEDON 1* (*CUC1*) and *CUC2* (Tian et al., 2020). The possibility that factors other than DNA sequence may contribute to REF6 binding affinity is supported by the fact that REF6 recognizes only 15% of the CTCTGYTY motifs in the Arabidopsis genome.

Protein–protein interaction between JMJ proteins and other transcription/chromatin factors are critical for H3K27me3 removal. However, conclusive *in vivo* evidence of when and where exactly those factors interact each other is lacking. Innovative *in vivo* imaging techniques are used to understand plant development and environmental responses through spatiotemporal regulation of gene expression (Abe et al., 2019; Hirakawa et al., 2019). Application of these techniques in H3K27me3 demethylase research to reveal the distribution of JMJ protein complexes will provide new insights into the spatiotemporal regulation of H3K27me3 removal.

CONCLUSION AND PERSPECTIVES

Flexible and robust gene expression during plant development and in response to the environment is primarily controlled by epigenetic regulation. In the past 5 years, plant epigenetic research on demethylases using transcriptome, epigenome, and crystal structure analyses has revealed the importance of H3K27me3. In *Arabidopsis thaliana*, the demethylation of H3K27 is regulated by a group of *JUMONJI DOMAIN-CONTANING PROTEIN* (*JMJ*) genes. *JMJ30* expression is high in various organs and is further boosted in response to environmental cues. On the other hand, the expression levels of *REF6*, *ELF6*, *JMJ13*, and *JMJ32* is moderate. *REF6* and *JMJ13* expression is also affected by environmental cues. These H3K27me3 demethylases bind to chromatin through generic or sequence-specific targeting mechanisms: direct binding to DNA via a ZnF domain, direct recognition of H3K27me3, or indirect binding through interactions with transcription factors. DNA methylation and minor groove width also fine-tune the binding affinity of these H3K27me3 demethylases. The targeting and occupancy of the histone demethylases on chromatin antagonize PRC2-mediated H3K27me3 deposition, and the removal of histone demethylases and prevention of their uncontrolled spread determine the shape of the H3K27me3 peak. Subsequently, H3K27me3 demethylases recruit a chromatin remodeler to activate gene transcription. One major limitation in current epigenome research is the scarcity of spatial information concerning the binding of epigenetic regulators and the nature of epigenetic modifications and co-factors. Furthermore, when and where target expression by H3K27me3 demethylases is mediated are poorly understood. Both binding patterns and DNA–protein structures and/or co-factors might vary among cells, tissues, and organs. In addition, growth conditions affect the epigenomic dynamics among individual plants. Therefore, specific genomic profiles obtained from plants grown under different conditions are needed to understand the specific roles of H3K27me3 demethylases during plant development and responses to the environment.

AUTHOR CONTRIBUTIONS

NY conceptualization and writing the manuscript.

FUNDING

This work was supported by a grant from the Japan Science and Technology Agency "PREST" (JPMJPR15QA), a JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (No. 18H04782), a JSPS KAKENHI Grant-in-Aid for Scientific Research B (No. 18H02465), a Grant-in-Aid for challenging Exploratory Research (No. 19K22431), and a grant from the SECOM Science and Technology Foundation to NY.

ACKNOWLEDGMENTS

The author thank Sachi Ando for critical comments on this manuscript.

REFERENCES

Abe, M., Kosaka, S., Shibuta, M., Nagata, K., Uemura, T., Nakano, A., et al. (2019). Transient activity of the florigen complex during the floral transition in *Arabidopsis thaliana*. *Development* 146:dev171504. doi: 10.1242/dev.171504

Antunez-Sanchez, J., Naish, M., Ramirez-Prado, J. S., Ohno, S., Huang, Y., Dawson, A., et al. (2020). A new role for histone demethylases in the maintenance of plant genome integrity. *Elife* 27:e58533. doi: 10.7554/eLife.58533

Bezhani, S., Winter, C., Hershman, S., Wanger, J. D., Kennedy, J. F., Kwan, C. S., et al. (2007). Unique, shared, and redundant roles for the Arabidopsis SWI/SNF chromatin remodeling ATPases BRHAMA and SPLAYED. *Plant Cell* 19, 403–416. doi: 10.1105/tpc.106.048272

Brand, U., Grunewald, M., Hobe, M., and Simon, R. (2002). Regulation of *CLV3* expression by two homeobox genes in *Arabidopsis*. *Plant Physiol.* 129, 565–575. doi: 10.1104/pp.001867

Bruneau, B. G., Koseki, H., Strome, S., and Torres-Padilla, M.-E. (2019). Chromatin and epigenetics in development: a Special Issue. *Development* 146:5025. doi: 10.1242/dev.185025

Chen, H., Tong, J., Fu, W., Liang, Z., Ruan, J., Yu, Y., et al. (2020). The H3K27me3 Demethylase RELATIVE OF EARLY FLOWERING6 Suppresses Seed Dormancy by Inducing Abscisic Acid Catabolism. *Plant Physiol.* 184, 1969–1978. doi: 10.1104/pp.20.01255

Crevillén, P. (2020). Histone demethylases as counterbalance to H3K27me3 silencing in plants. *iScience* 23:101715. doi: 10.1016/j.isci.2020.101715

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., et al. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature* 515, 587–590. doi: 10.1038/nature13722

Cui, X., Lu, F., Qiu, Q., Zhou, B., Gu, L., Zhang, S., et al. (2016). REF6 recognizes a specific DNA sequence to demethylate H3K27me3 and regulate organ boundary formation in Arabidopsis. *Nat. Genet.* 48, 694–699. doi: 10.1038/ng.3556

Eccleston, A., Cesari, F., and Skipper, M. (2013). Transcription and epigenetics. *Nature* 502:461. doi: 10.1038/502461a

Gan, E. S., Xu, Y., Wong, J. Y., Goh, J. G., Sun, B., Wee, W. Y., et al. (2014). Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of FLC in Arabidopsis. *Nat. Commun.* 5:5098. doi: 10.1038/ncomms6098

Hecker, A., Brand, L. H., Peter, S., Simoncello, N., Kilian, J., Harter, K., et al. (2015). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant Physiol.* 168, 1013–1024. doi: 10.1104/pp.15.00409

Hirakawa, T., Kuwata, K., Gallego, M. E., White, C. I., Nomoto, M., Tada, Y., et al. (2019). LSD1-LIKE1-Mediated H3K4me2 Demethylation Is Required for Homologous Recombination Repair. *Plant Physiol.* 181, 499–509. doi: 10.1104/pp.19.00530

Hou, X., Zhou, J., Liu, C., Liu, L., Shen, L., and Yu, H. (2014). Nuclear factor Y-mediated H3K27me3 demethylation of the SOC1 locus orchestrates flowering responses of Arabidopsis. *Nat. Commun.* 5:4601. doi: 10.1038/ncomms5601

Hyun, Y., Richter, R., Vincent, C., Martinez-Gallegos, R., Porri, A., and Coupland, G. (2016). Multi-layered regulation of SPL15 and cooperation with SOC1 integrate endogenous flowering pathways at the Arabidopsis shoot meristem. *Dev. Cell* 37, 1–13. doi: 10.1016/j.devcel.2016.04.001

Jones, M. A., Covington, M. F., DiTacchio, L., Vollmers, C., Panda, S., and Harmer, S. L. (2010). Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. *Proc. Natl. Acad. Sci. U. S. A* 107, 21623–21628. doi: 10.1073/pnas.1014204108

Keyzor, C., Mermaz, B., Trigazis, E., Jo, S., and Song, J. (2021). Histone demethylases ELF6 and JMJ13 antagonistically regulate self-fertility in Arabidopsis. *Front. Plant Sci.* 12:640135. doi: 10.3389/fpls.2021.640135

Lee, K., Park, O., and Seo, P. J. (2018). JMJ30-mediated H3K9me3 demethylation drives tissue identity changes to promote callus formation in Arabidopsis. *Plant J.* 95, 961–975. doi: 10.1111/tpj.14002

Li, C., Chen, C., Gao, L., Yang, S., Nguyen, V., Shi, X., et al. (2015). The *Arabidopsis* SWI2/SNF2 chromatin remodeler BRAHMA regulates Polycomb function during vegetative development and directly activates the flowering repressor gene *SVP*. *PLoS Genet* 11:e1004944. doi: 10.1371/journal.pgen.1004944

Li, C., Gu, L., Gao, L., Chen, C., Wei, C.-Q., Qiu, Q., et al. (2016). Concerted genomic targeting of H3K27 demethylase REF6 and chromatin remodeling ATPase BRM in Arabidopsis. *Nat. Genet.* 48, 687–693. doi: 10.1038/ng.3555

Liu, J., Feng, L., Gu, X., Deng, X., Qiu, Q., Li, Q., et al. (2019). An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in Arabidopsis. *Cell Res.* 29, 379–390. doi: 10.1038/s41422-019-0145-8

Lu, F., Cui, X., Zhang, S., Jenuwein, T., and Cao, X. (2011). Arabidopsis REF6 is a histone H3 lysine 27 demethylase. *Nat. Genet.* 43, 715–719. doi: 10.1038/ng.854

Lu, S. X., Knowles, S. M., Webb, C. J., Celaya, R. B., Cha, C., Siu, J. P., et al. (2011). The Jumonji C domain-containing protein JMJ30 regulates period length in the Arabidopsis circadian clock. *Plant Physiol.* 155, 906–915. doi: 10.1104/pp.110.167015

Lu, F., Li, G., Cui, X., Liu, C., Wang, X. J., and Cao, X. (2008). Comparative analysis of JmjC domain-containing proteins reveals the potential histone demethylases in Arabidopsis and rice. *J. Integr. Plant Biol.* 50, 886–896. doi: 10.1111/j.1744-7909.2008.00692.x

Lu, P., Porat, R., Nadeau, J. A., and O'Neill, S. D. (1996). Identification of a meristem L1 layer-specific gene in Arabidopsis that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *Plant Cell* 8, 2155–2168. doi: 10.1105/tpc.8.12.2155

Noh, B., Lee, S., Kim, H., Yi, G., Shin, E., Lee, M., et al. (2004). Divergent roles of a pair of homologous jumonji/zinc-finger-class transcription factor proteins in the regulation of Arabidopsis flowering time. *Plant Cell* 16, 2601–2613. doi: 10.1105/tpc.104.025353

Okamoto, M., Kuwahara, A., Seo, M., Kushiro, T., Asami, T., Hirai, N., et al. (2006). CYP707A1 and CYP707A2, which encode abscisic acid 8'-hydroxylases, are indispensable for proper control of seed dormancy and germination in Arabidopsis. *Plant Physiol.* 141, 97–107. doi: 10.1104/pp.106.079475.

Ouyang, W., Zhang, X., Peng, Y., Zhang, Q., Cao, Z., Li, G., et al. (2021). Rapid and Low-Input Profiling of Histone Marks in Plants Using Nucleus CUT&Tag. *Front Plant Sci.* 12, 634679. doi: 10.3389/fpls.2021.634679.

Qian, S., Wang, Y., Ma, H., and Zhang, L. (2015). Expansion and Functional Divergence of Jumonji C-Containing Histone Demethylases: Significance of Duplications in Ancestral Angiosperms and Vertebrates. *Plant Physiol.* 168, 1321–1337. doi: 10.1104/pp.15.00520

Qiu, Q., Mei, H., Deng, X., He, K., Wu, B., Yao, Q., et al. (2019). DNA methylation repels targeting of Arabidopsis REF6. *Nat. Commun.* 10:2063. doi: 10.1038/s41467-019-10026-1

Ryu, L. H., Huang, L., Kang, H. M., and Schiefelbein, J. (2019). Single-cell RNA sequencing resolves molecular relationships among individual plant cells. *Plant Physiol.* 179, 1444–1456. doi: 10.1104/pp.18.01482

Sakamoto, Y., Sato, M., Sato, Y., Harada, A., Suzuki, T., Goto, C., et al. (2020). Subnuclear gene positioning through lamina association affects copper tolerance. *Nat. Commun.* 11:5914. doi: 10.1038/s41467-020-19621-z

Tao, X., Feng, S., Zhao, T., and Guan, X. (2020). Efficient chromatin profiling of H3K4me3 modification in cotton using CUT&Tag. *Plant Methods* 16:120. doi: 10.1186/s13007-020-00664-8

Tian, C., Wang, Y., Yu, H., He, J., Wang, J., Shi, B., et al. (2019). A gene expression map of shoot domains reveals regulatory mechanisms. *Nat. Commun.* 10:141. doi: 10.1038/s41467-018-08083-z

Tian, Z., Li, X., Li, M., Wu, W., Zhang, M., Tang, C., et al. (2020). Crystal structures of REF6 and its complex with DNA reveal diverse recognition mechanisms. *Cell Discov.* 6:17. doi: 10.1038/s41421-020-0150-6

van Steensel, B., and Furlong, E. E. M. (2019). The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* 20, 327–337. doi: 10.1038/s41580-019-0114-6

Vergara, Z., and Gutierrez, C. (2017). Emerging roles of chromatin in the maintenance of genome organization and function in plants. *Genome Biol.* 18:96. doi: 10.1186/s13059-017-1236-9

Wang, X., Gao, J., Gao, S., Song, Y., Yang, Z., and Kuai, B. (2019). The H3K27me3 demethylase REF6 promotes leaf senescence through directly activating major senescence regulatory and functional genes in Arabidopsis. *PLoS Genet.* 15:1–24.

Winter, D., Venegar, B., Nahai, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2:e718. doi: 10.1371/journal.pone.0000718

Wu, J., Ichihashi, Y., Suzuki, T., Shibata, A., Shirasu, K., Yamaguchi, N., et al. (2019a). Abscisic acid-dependent histone demethylation during post-germination growth arrest in Arabidopsis. *Plant Cell Environ.* 42, 2198–2214. doi: 10.1371/journal.pgen.1008068

Wu, J., Yamaguchi, N., and Ito, T. (2019b). Histone demethylases control root elongation in response to stress-signaling hormone abscisic acid. *Plant Signal. Behav.* 14:1604019. doi: 10.1080/15592324.2019.1604019

Wu, J., Yan, M., Zhang, D., Zhou, D., Yamaguchi, N., and Ito, T. (2020). Histone demethylases coordinate the antagonistic interaction between abscisic acid and brassinosteroid signaling in Arabidopsis. *Front. Plant Sci.* 11:596835. doi: 10.3389/fpls.2020.596835

Wu, M. F., Sang, Y., Bezhani, S., Yamaguchi, N., Han, S. K., Li, Z., et al. (2012). SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3576–3581. doi: 10.1073/pnas.1113409109

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., et al. (2017). Cis and *trans* determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nat. Genet.* 49, 1546–1552. doi: 10.1038/ng.3937

Xiao, J., Lee, U. S., and Wagner, D. (2016). Tug of war: adding and removing histone lysine methylation in Arabidopsis. *Curr. Opin. Plant Biol.* 34, 41–53. doi: 10.1016/j.pbi.2016.08.002

Xiao, J., and Wagner, D. (2015). Polycomb repression in the regulation of growth and development in Arabidopsis. *Curr. Opin. Plant Biol.* 23, 15–24. doi: 10.1016/j.pbi.2014.10.003

Yamaguchi, N., Matsubara, S., Yoshimizu, K., Seki, M., Hamada, K., Kamitani, M., et al. (2020) H3K27me3 demethylases alter *HSP22* and *HSP17.6C* expression in response to recurring heat in Arabidopsis. *Nat. Commun.* (in press). doi: 10.1038/s41467-021-23766-w

Yan, W., Chen, D., Smaczniak, C., Engelhorn, J., Liu, H., Yang, W., et al. (2018). Dynamic and spatial restriction of Polycomb activity by plant histone demethylases. *Nat. Plants* 4, 681–689. doi: 10.1038/s41477-018-0219-5

Yan, Y., Shen, L., Chen, Y., Bao, S., Thong, Z., and Yu, H. (2014). A MYB-domain protein EFM mediates flowering responses to environmental cues in Arabidopsis. *Dev. Cell* 30, 437–448. doi: 10.1016/j.devcel.2014.07.004

Yang, H., Howard, M., and Dean, C. (2016). Physical coupling of activation and derepression activities to maintain an active transcriptional state at FLC. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9369–9374. doi: 10.1073/pnas.1605733113

Yu, X., Li, L., Li, L., Guo, M., Chory, J., and Yin, Y. (2008). Modulation of brassinosteroid-regulated gene expression by Jumonji domain-containing proteins ELF6 and REF6 in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7618–7623. doi: 10.1073/pnas.0802254105

Zheng, S., Hu, H., Ren, H., Yang, Z., Qiu, Q., Qi, W., et al. (2019). The Arabidopsis H3K27me3 demethylase JUMONJI 13 is a temperature and photoperiod dependent flowering repressor. *Nat. Commun.* 10:1303. doi: 10.1038/s41467-019-09310-x

Zheng, X. Y., and Gehring, M. (2019). Low-input chromatin profiling in Arabidopsis endosperm using CUT&RUN. *Plant Reprod.* 32, 63–75. doi: 10.1007/s00497-018-00358-1

Zhou, Y., Wang, Y., Krause, K., Yang, T., Dongus, J. A., Zhang, Y., et al. (2018). Telobox motifs recruit CLF/SWN–PRC2 for H3K27me3 deposition via TRB factors in Arabidopsis. *Nat. Genet.* 50:638. doi: 10.1038/s41588-018-0109-9

Zhu, Y., Klasfeld, S., Jeong, C. W., Jin, R., Goto, K., Yamaguchi, N., et al. (2020). TERMINAL FLOWER1-FD complex target genes and competition with FLOWERING LOCUS T. *Nat. Commun.* 11:5118. doi: 10.1038/s41467-020-18782-1

Conflict of Interest: The author declares 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 in Plant Science

ORIGINAL RESEARCH
published: 01 July 2021
doi: 10.3389/fpls.2021.703994

Comprehensive Analysis of the SBP Family in Blueberry and Their Regulatory Mechanism Controlling Chlorophyll Accumulation

Xin Xie[1†], Shaokang Yue[1†], Baosheng Shi[2†], Hongxue Li[1], Yuhai Cui[3,4], Jingying Wang[1], Pengjie Yang[1], Shuchun Li[5], Xuyan Li[1] and Shaomin Bian[1]**

[1] College of Plant Science, Jilin University, Changchun, China, [2] College of Landscape Architecture and Tourism, Hebei Agricultural University, Baoding, China, [3] London Research and Development Centre, Agriculture and Agri-Food Canada, London, ON Canada, [4] Department of Biology, Western University, London, ON, Canada, [5] Department of Pain, Second Hospital of Jilin University, Changchun, China

OPEN ACCESS

Edited by:
Gang Wu,
Zhejiang Agriculture and Forestry University, China

Reviewed by:
Jiaqiang Sun,
Chinese Academy of Agricultural Sciences (CAAS), China
Jiyuan Shen,
South China Agricultural University, China

***Correspondence:**
Shaomin Bian
shmbian@jlu.edu.cn
Xuyan Li
xuyanli@jlu.edu.cn

[†] These authors have contributed equally to this work

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

Received: 01 May 2021
Accepted: 09 June 2021
Published: 01 July 2021

Citation:
Xie X, Yue S, Shi B, Li H, Cui Y, Wang J, Yang P, Li S, Li X and Bian S (2021) Comprehensive Analysis of the SBP Family in Blueberry and Their Regulatory Mechanism Controlling Chlorophyll Accumulation. Front. Plant Sci. 12:703994. doi: 10.3389/fpls.2021.703994

SQUAMOSA Promoter Binding Protein (SBP) family genes act as central players to regulate plant growth and development with functional redundancy and specificity. Addressing the diversity of the SBP family in crops is of great significance to precisely utilize them to improve agronomic traits. Blueberry is an important economic berry crop. However, the SBP family has not been described in blueberry. In the present study, twenty *VcSBP* genes were identified through data mining against blueberry transcriptome databases. These VcSBPs could be clustered into eight groups, and the gene structures and motif compositions are divergent among the groups and similar within each group. The VcSBPs were differentially expressed in various tissues. Intriguingly, 10 VcSBPs were highly expressed at green fruit stages and dramatically decreased at the onset of fruit ripening, implying that they are important regulators during early fruit development. Computational analysis showed that 10 *VcSBP*s were targeted by miR156, and four of them were further verified by degradome sequencing. Moreover, their functional diversity was studied in Arabidopsis. Noticeably, three VcSBPs significantly increased chlorophyll accumulation, and qRT-PCR analysis indicated that VcSBP13a in Arabidopsis enhanced the expression of chlorophyll biosynthetic genes such as AtDVR, AtPORA, AtPORB, AtPORC, and AtCAO. Finally, the targets of VcSBPs were computationally identified in blueberry, and the Y1H assay showed that VcSBP13a could physically bind to the promoter region of the chlorophyll-associated gene *VcLHCB1*. Our findings provided an overall framework for individually understanding the characteristics and functions of the SBP family in blueberry.

Keywords: blueberry, SBP gene, miR156, chlorophyll accumulation, SBP targets

INTRODUCTION

Blueberry (*Vaccinium* spp.) is a globally cultivated perennial shrub with outstanding economic value. Its fruit is not only sweet but also rich in nutrients, especially anthocyanins, which greatly promote human health such as improvement of vision, blood glucose balance, elimination of free radicals, aging delay, inhibition of obesity and hyperlipidemia, and prevention of cardiovascular diseases (Routray and Orsat, 2011). Thus, blueberry growth and development, especially the events related to fruit ripening

and anthocyanin biosynthesis, have started to attract attention in recent years. To date, a few regulators have been shown to be involved in the regulation of blueberry growth and development, including transcription factor genes *VcMYB*s, *VcSOC1-k*, *VcDDF1*, *VcFT*, some miRNAs, and hormones (IAA and ABA) (Zifkin et al., 2012; Song et al., 2013; Walworth et al., 2016; Hou et al., 2017, 2020; Song and Gao, 2017; Plunkett et al., 2018). Recently, high-throughput sequencing data provided considerable information for identifying and characterizing the regulators that control blueberry growth and development (Rowland et al., 2012; Hou et al., 2017; Qi et al., 2019). However, our understanding of the regulatory network underlying blueberry growth and development are extremely limited.

SQUAMOSA Promoter Binding Proteins (SBPs) constitute a plant-specific transcription factor family featured by a highly conserved SBP domain of 76 amino acids. Generally, the SBP domain harbors three common structures: two tandem zinc fingers (C3H and C2HC) and a nuclear localization signal (NLS), partially overlapping with the second zinc finger at the C-terminal (Birkenbihl et al., 2005). It has been well known that SBP proteins can bind to a consensus DNA sequence TNCGTACAA with the GTAC as the binding core, therefore regulating the expression of their target clients (Birkenbihl et al., 2005; Kropat et al., 2005). SBP proteins play important roles in various biological and cellular processes through regulating their target clients, spanning virtually every aspect of plant growth and development as well as stress response. These include leaf morphology and leaf initiation (Preston et al., 2016), trichome formation (Yu et al., 2010), phase transition (Xu et al., 2016), shoot branching and maturation (Gao et al., 2018), regeneration of shoot and root (Barrera-Rojas et al., 2020; Ye et al., 2020), root development (Yu et al., 2015), flowering (Xie et al., 2020), male fertility (Xing et al., 2010), ovary and fruit development (Silva et al., 2014), cell number and size (Usami et al., 2009), and grain yield (Wang et al., 2017), etc. Evidently, *SBP* genes are a class of central players in the regulation of plant growth and development, which can be utilized for the improvement of important agronomic traits.

In 1996, the first two *SBP*s were identified in *Antirrhinum majus*, and shown to regulate the expression of the MADS-box gene *SQUAMOSA* directly through binding to its promoter region, therefore controlling flowering (Klein et al., 1996). With the availability of whole-genome information and transcriptome data, *SBP* genes have been isolated in many plant species, from the model plant Arabidopsis to economically important crops (Salinas et al., 2012; Hou et al., 2013; Bhogale et al., 2014; Li and Lu, 2014; Shalom et al., 2015). The SBP family is a relatively small group of transcription factors in plants, and the SBP family members show diverse features and evolutionary divergences. Emerging evidence indicated that SBPs exert their regulatory functions in a member-specific manner. For example, *SBP-like 9* (*SPL9*) in *Arabidopsis thaliana* might serve as a negative regulator of wall ingrowth deposition in transfer cells of phloem parenchyma (Nguyen et al., 2017), whereas *SPL3* cannot affect the deposition of wall ingrowth but enhance phosphate-deficient response (Lei et al., 2016). Likewise, *OsSPL14* acts in controlling rice tillering growth (Luo et al., 2012), and *OsSPL16* was found to be a regulator of grain size, shape, and quality in *Oryza sativa* (Wang et al., 2012). Nevertheless, a number of studies showed that members of the SBP family could be functionally redundant in the regulation of plant growth and development. For example, *AtSPL3/4/5* redundantly promote flowering through activating the expression of *LEAFY*, *FRUITFULL*, and *APETALA1* (Jung et al., 2016), while *AtSPL9/15* and *AtSPL2/10/11* act as regulators of plastochron and branching (Schwarz et al., 2008; Shikata et al., 2009). Additionally, a subset of *SBP* genes can be subjected to miR156-guided transcriptional cleavage and translational repression, for example, 11 out of the 17 *SPL*s in Arabidopsis and seven out of the 19 *SPL*s in pear, thereby being integrated into miR156/*SPL* modules to regulate plant growth, development, and stress response (Zhang et al., 2015; Qian et al., 2017). Clearly, the SBP family members show distinct features and perform their functions with redundancy and specificity. Thus, addressing the diversity and specificity of the SBP family in different crop species is of great significance in order to precisely utilize them to improve agronomic traits.

It has been accepted that the functional roles of *SBP* genes are highly conserved across plant species. However, novel functions of *SBP* genes have been constantly revealed in crop species with special developmental processes or organs (Bhogale et al., 2014; Silva et al., 2014; Qian et al., 2017). Fruit growth and ripening is a specific process for fruit-bearing plant species, and many fruit-specific events occur during the process. Accumulating evidence indicates that *SBP* genes are involved in the regulation of fruit growth and ripening. For instance, the Colorless non-ripening (CNR) locus of tomato (a homolog of *AtSPL3*) is crucial for fruit ripening (Manning et al., 2006), while *SPL18* in grape might regulate berry development at the veraison stage in an ABA-independent manner (Xie et al., 2019). Likewise, *VmTDR4* (a *SQUAMOSA*-class *MADS-box* gene) is positively involved in the regulation of anthocyanin accumulation during bilberry fruit ripening (Jaakola et al., 2010), while MaSPL16 in banana regulates carotenoid biosynthesis through promoting the expression of *MaLCYB*s genes (Zhu et al., 2020). These functions were not observed in non-fleshy-fruited plant species such as Arabidopsis and rice. Obviously, it is of great interest to comprehensively characterize the SBP family in crops with special developmental processes or organs and to reveal their functional roles and neo-functionalization.

Since the SBP family proteins are powerful regulators with functional diversification in plants, study of these genes will enhance understanding of the regulatory network underlying blueberry growth and development. To date, however, the characteristics and functional diversity of the SBP family have remained unexplored in blueberry. In recent years, the transcriptional profiles of blueberry leaves, flower buds, and fruits at different development stages have been investigated using high-throughput sequencing technology (Rowland et al., 2012; Gupta et al., 2015; Li et al., 2016). These transcriptome data have enabled the identification of the *SBP* family genes involved in blueberry growth and development. In the present study, 20 *VcSBP*s were identified from the blueberry transcriptome database. Gene structure, phylogeny, motif composition, miRNA

target sites, and expression patterns in different tissues were systematically analyzed. Furthermore, the functional diversity of the *VcSBP* family genes were studied in Arabidopsis. Additionally, the targets of VcSBP proteins were investigated in blueberry. These findings lay a foundation for further studying the functional roles of the *SBP* genes and their regulatory mechanisms during blueberry growth and development, which will contribute to the improvement of blueberry agronomic traits.

MATERIALS AND METHODS

Plant Materials

Seven-year-old blueberry trees (*Vaccinium corymbosum*, cv. Northland) from clonal propagation were grown at the experimental station at Jilin University (Changchun, China). Blueberry tissues were randomly harvested from six different seven-year-old blueberry plants, including new leaf, young shoot, unopened flower, opening flower, and fruit at six developmental stages [green pad (FS1), green cup I (FS2), green cup II (FS3), light green/white (FWS), pink (FPS) and blue (FMS) fruits] (Li et al., 2020), frozen in liquid nitrogen and stored at −80°C.

Arabidopsis and tobacco (*Nicotiana benthamiana*) plants were grown in growth chambers under long-days (16 h light/8 h dark) at 20°C with 70–80% relative humidity.

Identification of *SBP* Genes in Blueberry

The CDS sequences of *SBP* genes from Arabidopsis and grape were downloaded from the publicly available databases TAIR[1] or Phytozome[2], and then used as reference sequences to perform local blast searches for querying their homologs against the publicly available transcriptome databases of blueberry[3] and our previously assembled transcriptome data. The conserved SBP-specific domains were confirmed using the PROSITE Server[4], and all of the *SBP-like* genes without an SBP domain were discarded. The physicochemical properties, including molecular weight (MW), and isoelectric point (pI), of the identified SBP proteins, were predicted using the ExPASy Compute pI/Mw tool[5].

Chromosomal Location and Phylogenetic Analysis of the *VcSBP* Family Genes

All *VcSBP* genes were mapped to the genome of *V. corymbosum*, cv. Draper, according to the approximate location information (Colle et al., 2019), and their positions were imported into the CIRCOS software to generate a circle plot (Krzywinski et al., 2009). The SBP protein sequences (17 from grape, 27 from apple, and 17 from tomato) were downloaded from Phytozome (see text footnote 2). All the SBP protein sequences from blueberry, Arabidopsis, grape, apple, and tomato were used for phylogenetic analysis, and phylogenetic trees were constructed with the MEGA7.0 software using the maximum likelihood with 1000 bootstrap replications (Kumar et al., 2016). The sequence logo was created using Weblogo online software[6].

Analysis of Gene Structure and Conserved Protein Motifs

The exon/intron structure of each *VcSBP* gene was analyzed using the Gene Structure Display Server[7] by comparing the coding sequence and genomic sequence. Potentially conserved motifs of VcSBP proteins were predicted using the online Multiple Expectation Maximization for Motif Elucidation (MEME) toolkit[8], with the following parameter settings: the minimum motif width = 20, the maximum motif width = 50, and the maximum number of motifs = 20.

MicroRNA Target Prediction

To identify *VcSBP*s targeted by miR156/157, the coding regions and 3′ UTRs of all *VcSBP* sequences were analyzed at the psRNATarget server[9] with blueberry miR156/157 mature sequences (Hou et al., 2017). The sequence logo of miR156/157 was created using the Weblogo online software (see text footnote 6).

Expression *P*attern *A*nalysis of VcSBP *G*enes in *B*lueberry

Total RNAs were isolated from blueberry leaf, shoot, unopened flower, opening flower, fruit tissues at six developmental stages, and blueberry tissue culture seedlings as well as Arabidopsis leaf. First-strand cDNA was synthesized using the PrimeScript™ RT reagent kit with gDNA Eraser (Takara, Japan). qRT-PCR was subsequently conducted with an ABI StepOnePlus PCR system and SYBR Premix Ex Taq (Takara, Japan). Blueberry *ACTIN* was set as an internal reference for data normalization. Three biological replicates with three technical replicates were performed for each sample, and data were analyzed by the software ABI StepOnePlus v2.3 and one-way ANOVA with LSD test, and p-value < 0.05 was considered to be statistically significant. Primer information is listed in **Supplementary Table 1**.

Vector Construction and Plant Transformation

The full-length CDS of each *VcSBP* was amplified using gene-specific primers (**Supplementary Table 1**). All purified PCR products were cloned into the Gateway entry vector pDONR207 and then transferred into the destination vector pEarleyGate101 (pEG101) through homologous recombination. All the constructed plasmids were confirmed by PCR and sequencing. The expression vectors (pEG101-*VcSBP*s) were individually transformed into *Agrobacterium tumefaciens* strain GV3101.

Arabidopsis transformation was conducted using the floral dip method described by Zhang et al. (2006). Transgenic lines

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

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

[3]www.vaccinium.org

[4]https://prosite.expasy.org/

[5]https://web.expasy.org/compute_pi/

[6]http://weblogo.threeplusone.com/

[7]https://gsds.cbi.pku.edu.cn/

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

[9]http://plantgrn.noble.org/psRNATarget/

were screened in the soil with 200 μg/mL glufosinate and then confirmed by PCR with gene-specific primers. Primer information is listed in **Supplementary Table 1**.

The *VcMIR156a* gene was constructed into pBI121 as described previously (Li et al., 2020) and transferred into *Agrobacterium tumefaciens* strain EHA105. Blueberry transformation was performed according to the method described by Song and Sink (2006). The transgenic blueberry lines were obtained and confirmed by PCR with gene-specific primers (**Supplementary Table 1**). The *VcMIR156a*-overexpressing transgenic Arabidopsis were generated as previously described (Li et al., 2020).

Prediction of *VcSBP* Targets in Blueberry

The genes containing the TNCGTACAA element within 2000 bp upstream were extracted against the reference genome of blueberry (*V. corymbosum*, cv. Draper) (Colle et al., 2019). To functionally annotate these targets, all protein sequences were analyzed using eggNOG-Mapper[10]. Density distribution of distance was visualized using ggplot2 in R.

Yeast One-Hybrid (Y1H) Assay

To investigate the interaction of VcSBPs and their targets, the full-length CDS of *VcSBP13a* was cloned and introduced into the vector pB42AD. The fragment 597-796 bp upstream of the transcription start site (TSS) of *VcLHCB1* containing two TNCGTACAA elements and the fragment 1096–1548 upstream of the TSS of *VcLHCB2* with five GTAC elements were cloned as promoter regions (*pVcLHCB1* and *pVcLHCB2*) and constructed into the vector pLacZi, respectively. Three negative controls, i.e., *pB42AD/pLacZi*, *pB42AD-VcSBP13a/pLacZi*, and *pB42AD/pLacZi-pVcLHCB1*, *pB42AD/pLacZi-pVcLHCB2*, and one positive control, *pB42AD-AtRVE8/placZi-AtPRR5*, were also generated. Different plasmid combinations were separately co-transformed into the yeast cells (EGY48). The primers are listed in **Supplementary Table 1**.

RESULTS

Identification of *SBP* Genes and Their Characterization in Blueberry

To identify *SBP* genes in blueberry, the CDS sequences of the *SBP* genes from both *Vitis vinifera* and Arabidopsis were used as queries to conduct BLASTn against the *V. corymbosum* GDV RefTrans V1 and our previously assembled transcriptome data (Hou et al., 2017). After removal of redundant sequences, a total of 22 *SBP* sequences were identified in blueberry, which are then named as *VcSBP* and each of them assigned a species number corresponding to their closest homolog in Arabidopsis (**Supplementary Figure 1**). To verify the sequences of the *VcSBP* genes, their full-length CDSs were amplified and sequenced, and the results showed that all the cloned *SBP* genes are indeed the same sequence as listed in the Genome Database for *V. corymbosum* cv. Draper v1.0.

[10]http://eggnog-mapper.embl.de/

The features of all the *VcSBP* family members were computationally characterized. As shown in **Supplementary Table 2**, the CDS lengths of these *VcSBP* genes are quite variable, ranging from 363 to 3222 bp, which is consistent with the *SBP* family in other plant species such as Arabidopsis, apple, grape (Hou et al., 2013; Li et al., 2013; Zhang et al., 2015). Their deduced proteins were estimated to possess the theoretical pI values from 5.96 to 10.54 and the MWs from 23.17 to 117.97 kDa. Furthermore, the SBP domains were analyzed using the online tool CD search[11]. As shown in **Figures 1A,B**, all the SBP proteins, except VcSBP6c, VcSBP14aAS, and VcSBP14cAS, contain a typical SBP domain featured by two zinc finger structures (C3H and C2HC) and an NLS motif. VcSBP14aAS harbors an SBP domain with the absence of C3H and an incomplete C2HC, while the SBP domains in VcSBP6c and VcSBP14cAS lack the NLS motif (**Figures 1A,B**). These results indicated that the 22 putative genes are SBP family members.

Distribution of *VcSBP* Genes in the Blueberry Genome and Their Evolutionary Relationships

To date, the draft genome assembly of *V. corymbosum* contains 1760 scaffolds[12]. To map the locations of the *VcSBP* family genes in the draft genome, a Circos map was generated using the corresponding scaffolds where the *VcSBP* genes are situated. It turns out that they are unevenly distributed in 15 different scaffolds (i.e., 1, 2, 4, 5, 8, 11, 12, 22, 23, 24, 25, 26, 34, 35, 883, **Figure 1C**). Further observation indicated that two pairs of *VcSBP* genes (*VcSBP14a* and *VcSBP14aAS*, *VcSBP14c*, and *VcSBP14cAS*) were situated at the same loci with the similarity of 65.01 and 93.09%, respectively, indicating that they might be derived from different transcript splicing of the same genes. Thus, the 22 *SBP* sequences were likely derived from 20 *SBP* genes and two alternative splices. Gene family expansion can arise from gene duplication events such as tandem duplication and segmental duplication of chromosomal regions (Leister, 2004). Generally, tandem duplication refer to those closely related genes separated by the distance within 50 kb in the same chromosome. It was observed that the distance between *VcSBP8a* and *VcSBP8b* is 15,110 bp (**Figure 1C** and **Supplementary Table 2**), suggesting that they might be derived from tandem duplication. In contrast, the distances between *VcSBP7a* and *VcSBP7b* as well as *VcSBP14b* and *VcSBP14c* are relatively far from each other, and their similarities reach 99.27 and 94.26%, respectively, implying that they were possibly generated from segmental duplication.

To explore the evolutionary relationships among the SBP family proteins, a phylogenetic tree was generated using the protein sequences of VcSBPs and the SBPs from apple, grape, tomato, and Arabidopsis. As shown in **Supplementary Figure 2**, all the SBP proteins were classified into six different groups (G1-G6), and VcSBPs were separately distributed to the 6 groups, suggesting that the VcSBP family might

[11]www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

[12]https://www.vaccinium.org/analysis/49

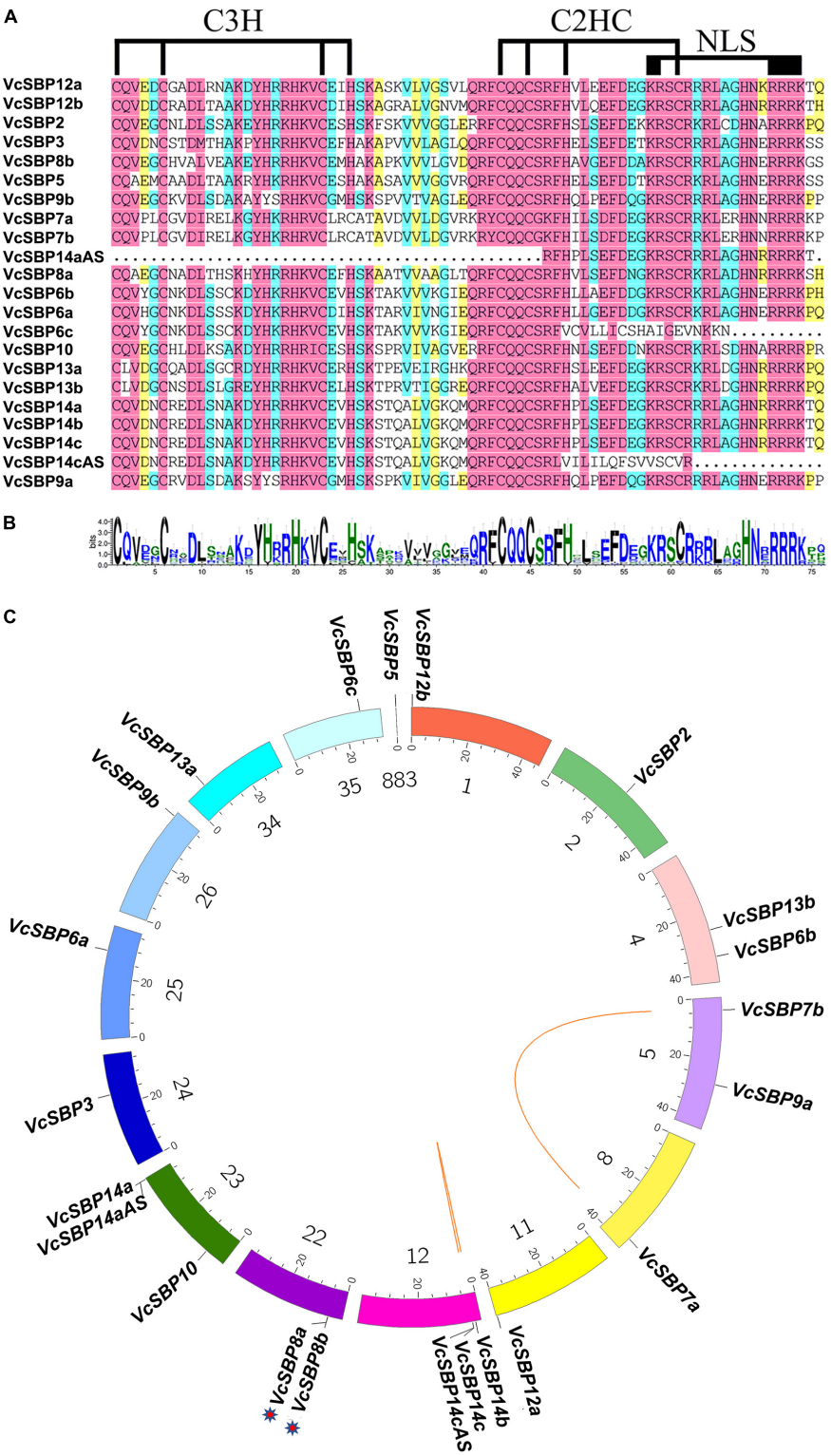

FIGURE 1 | The SBP domains and chromosomal localization of the *VcSBP* family genes. **(A)** Multiple alignment of the SBP domains. The two conserved zinc-finger structures (C3H and C2HC) and the NLS are indicated. **(B)** Sequence logo of the SBP domains in VcSBPs. The total height of each stack represents the conservation degree of each position, while the height of the letters within each stack indicates the relative frequency of the corresponding amino acid. **(C)** Chromosomal localization and duplication of *SBP* genes in blueberry. Each colored box represents a scaffold. The approximate distribution of each *VcSBP* gene is marked on the circle with a short black line. The tandem duplication cluster is indicated with stars. Colored lines indicate the linkage group with segmental duplication.

have experienced evolutionary diversification similar to those in the other four plant species. For example, seven small VcSBP proteins with no more than 254 aa (VcSBP3, VcSBP5, VcSBP6c, VcSBP9a, VcSBP10, VcSBP14aAS, and VcSBP14cAS) were separately distributed into the six groups, while the large proteins with more than 800 aa (VcSBP14a, VcSBP12a, and VcSPB7a/7b) were clustered into G5 and G6, respectively. Further observation indicated that VcSBPs were closer to their homologs from apple, grape, Arabidopsis and/or tomato in the phylogenetic tree. For instance, VcSBP3 was grouped together with SlySBP3, AtSBP3, VvSBP9, CNR, and while VcSBP2 was distributed into the subgroup of SBP2/10/11 with the inclusion of VvSBP2, SlySBP2, and AtSBP2.

VcSBP Family Shows Diverse Gene Structures and Motif Compositions

To understand the structural diversity of *VcSBP* family genes, the exon/intron structures were generated according to the gene coding and genomic sequences. Consistent with previous reports in other plant species (Hou et al., 2013; Li et al., 2013; Zhang et al., 2015), *VcSBP* genes showed a high variation in the number of exons. As indicated in **Figure 2A**, four *VcSBP* genes (*VcSBP7a*, *VcSBP7b*, *VcSBP12a*, and *VcSBP14a*) comprise 10 exons with intron intervals. In contrast, *VcSBP6c*, *VcSBP14aAS*, and *VcSBP14cAS* harbor only one exon without intron. The remaining *VcSBP*s have 2–4 exons. Furthermore, integration analysis of exon/intron structures with phylogenetic relationship and sequence identity was conducted. It turns out that the pairs of *VcSBP*s in the same clade basically display similar exon/intron structures (**Figure 2A**). Two pairs of duplicated genes (*VcSBP7a* and *VcSBP7b*, *VcSBP14b*, and *VcSBP14c*) show not only similar exon/intron structure but also high similarity with the values of 99.27 and 94.21%, respectively (**Supplementary Table 2**), supporting that they might undergo similar exon/intron gain or loss events with less functional diversification. However, the remaining *VcSBP* pairs with similar exon/intron structure in the same clade displayed relatively low similarities ranging from 8–48% (**Figure 2A** and **Supplementary Table 2**), implying the diversity of their functional roles.

To provide clues about the functional diversity of VcSBP family, conserved motifs in each of the VcSBP proteins were predicted using the online tool ScanProsite. As shown in **Figure 2B**, twenty conserved motifs were identified in VcSBPs, and two motifs (the motifs 1 and 2) constitute the SBP domain. Five VcSBPs (VcSBP7a/7b and VcSBP14a/14b/14c) harbor 11–13 motifs, while the remaining VcSBPs contain 2–5 motifs. Although most of the 20 motifs are functionally unknown, the existence of multiple motif compositions implied the functional diversity of the SBP family in blueberry. The 20 VcSBPs and two alternatively spliced species were clustered into eight groups in the phylogenetic tree. It was observed that the VcSBP proteins in the same group in the phylogenetic tree basically show similar motif composition, suggesting possible functional redundancy within the same group.

*VcSBP*s Are Differentially Expressed in Different Tissues and Throughout Fruit Development

To obtain clues about the functional roles of *VcSBP* genes, their expression patterns in five tissues (new leaf, young shoot, opening, and unopened flower, and mature fruit; **Figure 3A**) were examined using qRT-PCR. Since high sequence similarity exists within each of the three *VcSBP* groups (*VcSBP7a/b*, *VcSBP6b/c*, and *VcSBP14a/b/c/cAS*), only one gene was chosen as representative for each group (*VcSBP6b*, *VcSBP7a*, and *VcSBP14a*). The examined *VcSBP* genes were found to be differentially expressed in the five tissues. As shown in **Figures 3B,C**, 10 *SBP* genes showed the highest expression in shoot, especially *VcSBP13b* and *VcSBP9a*, with 7.46–509.52 and 7.12–118.31-fold increase as compared to the other four tissues. Meanwhile, three *VcSBP* genes (*VcSBP8a*, *VcSBP8b*, and *VcSBP12a*) were highly expressed in opening flower (**Figure 3D**), and three *VcSBP* genes (*VcSBP5*, *VcSBP6b*, and *VcSBP13a*) in unopened flower and shoot (**Figure 3E**). In mature fruit, all the *VcSBP* genes were expressed at relatively low levels except *VcSBP9b*, *VcSBP12b*, *VcSBP14a*, and *VcSBP14aAS* (**Figure 3F**). These results suggested that the *VcSBP* family might perform functions in an organ-specific manner.

Blueberry fruit development can be generally divided into three phases: fruit growth, a transition from growth to maturation, and maturation (Zifkin et al., 2012). To explore the functional roles of *SBP* family during fruit development, the expression patterns of *VcSBP*s were investigated in fruits at six developmental stages (green pad, green cup I, green cup II, light green/white, pink, and blue fruit, **Figure 3A**). The three early developmental stages represent the growth phase; the light green/white stage corresponds to the transition stage; the pink and blue stages refer to the maturation phase. As shown in **Figures 3G,H**, 10 *VcSBP* genes were highly expressed at the three early developmental stages (green pad, green cup I, green cup II), and dramatically decreased at the light green stage (especially *VcSBP3*, *VcSBP5*, *VcSBP9b*, *VcSBP10*, and *VcSBP13b* with more than 10-fold changes as compared to the ones at green cup II), and then remained at a low level until fruit maturation. Conversely, the expression levels of some *VcSBP* genes (*VcSBP9b*, *VcSBP12a*, *VcSBP12b*, *VcSBP14a*, and *VcSBP14aAS*) were relatively low at the three early developmental stages, but then increased from the light green stage until fruit maturation (**Figure 3I**). Also, it was observed that *VcSBP8b* was gradually increased from the green cup I stage to the light green stage, and then remarkably decreased at the maturation stage (**Figure 3J**). These results suggested that *VcSBP* family might play different, even opposite, roles during blueberry fruit development.

A Subset of *SBP* Genes Are Targeted by miR156 in Blueberry

It is well acknowledged that most SBP family members can be regulated through miR156/157-mediated mRNA cleavage or translational repression in plants (Wang and Wang, 2015). Previously we identified six *MIR156/MIR157* genes in blueberry

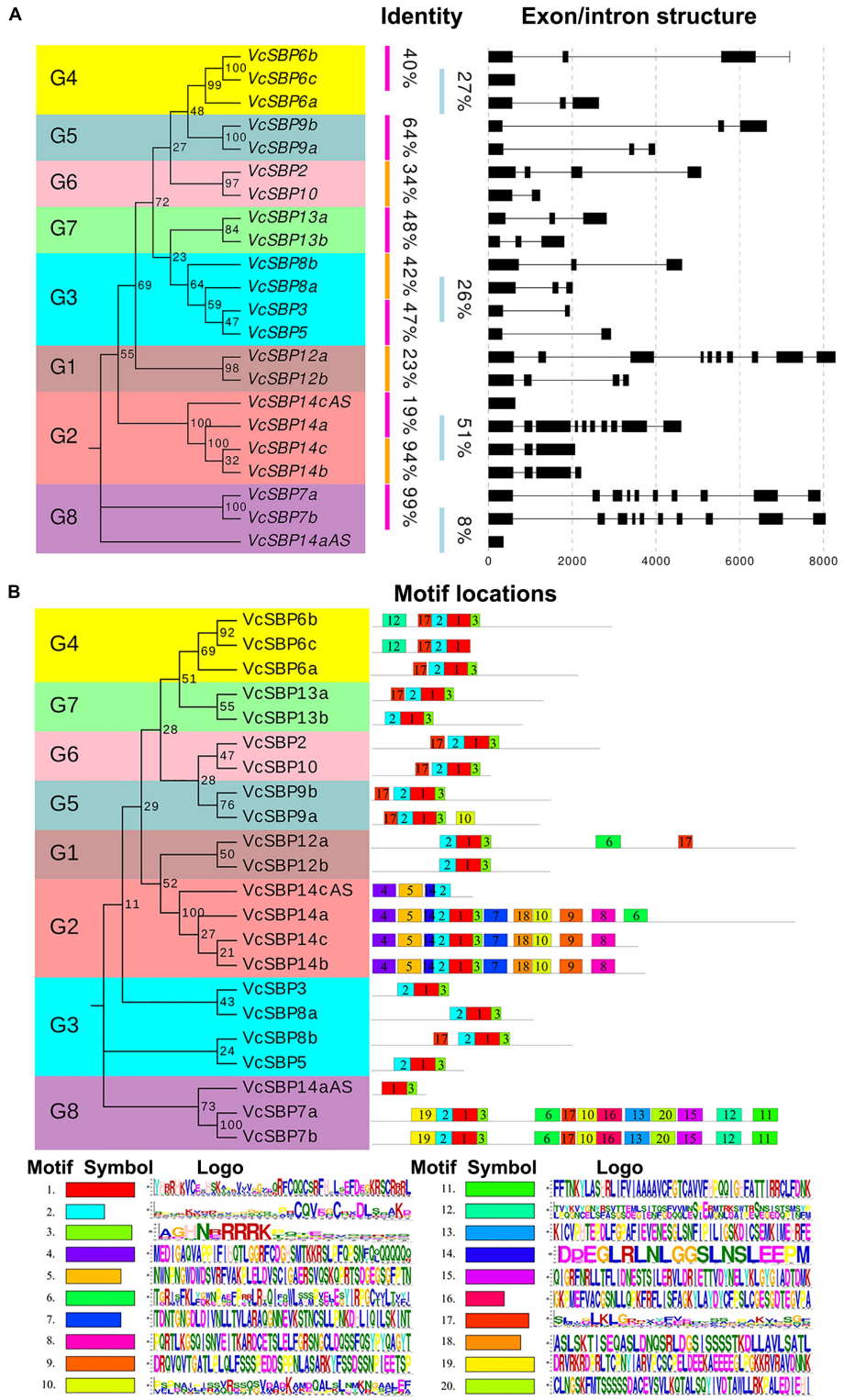

FIGURE 2 | Gene structures and protein motif compositions of the VcSBP family. **(A)** The exon/intron structures of *VcSBP* genes. The left panel is the phylogenetic tree of *VcSBP* genes. Eight groups are clustered (G1–G8), and the percent similarity between the gene pair is listed. The right panel shows the intron-exon structures where the exons are shown by rectangular, and the introns are represented by thin lines. **(B)** Motif analysis of VcSBP proteins. The left panel is the phylogenetic tree of VcSBP proteins, and eight groups are clustered (G1–G8). The right panel shows the motif compositions of VcSBP proteins. The motifs were identified using the program MEME, represented with boxes of different colors labeled by number (1–20). The sequence logos of 20 motifs are listed, and the height of the letters within each stack indicates the relative frequency. The symbol and the ID number are corresponding to the colored box and the number in the right panel.

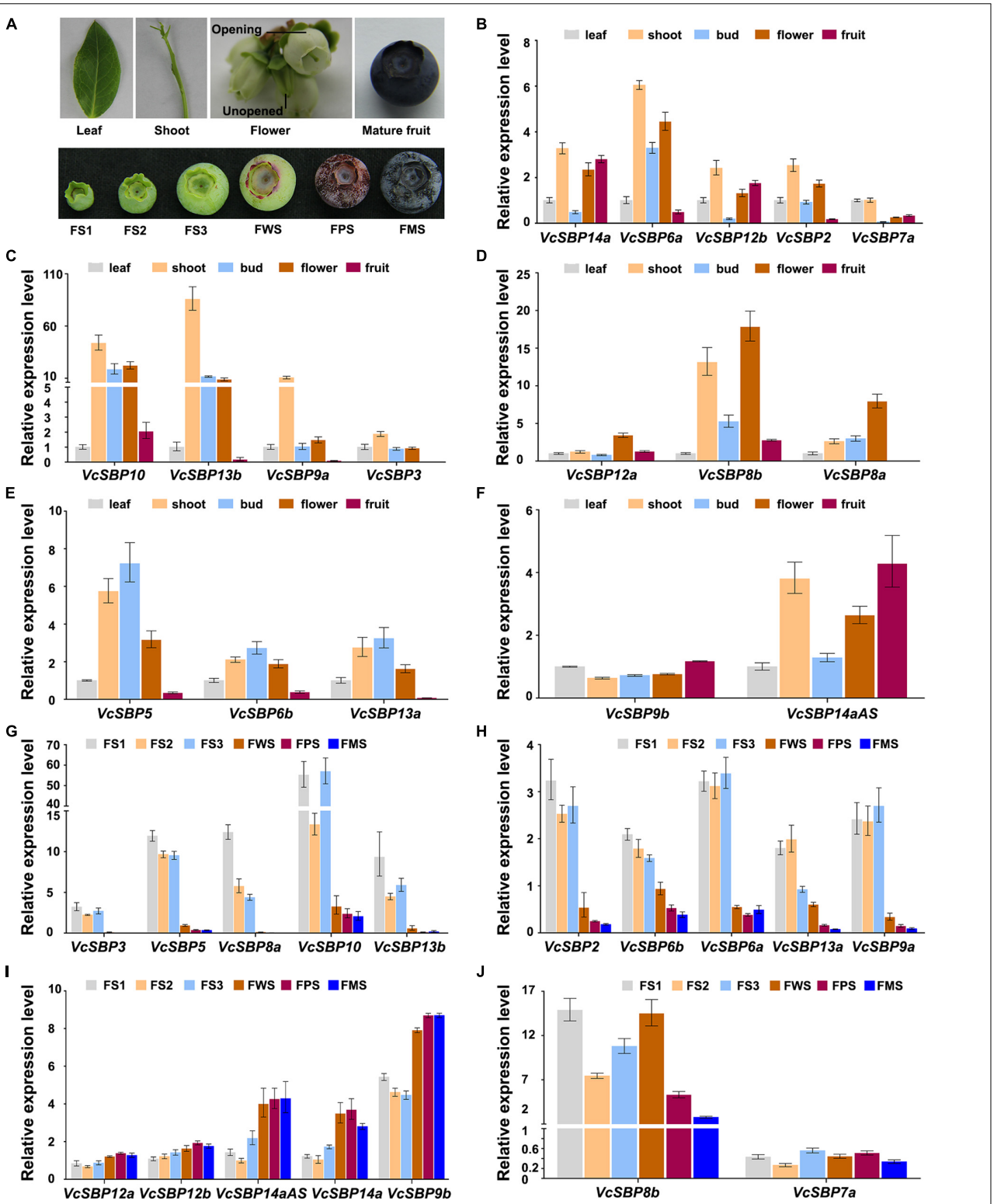

FIGURE 3 | Expression patterns of *VcSBP* genes in different tissues and during fruit development. **(A)** Photograph of different tissues and fruits at six developmental stages. Different tissues include blueberry new leaf, young shoot, unopened flower, opening flower, and mature fruit, while FS1, FS2, FS3, FWS, FPS, and FMS refer to green pad, green cup I, green cup II, light green/white, pink, and blue stages, respectively. **(B–F)** Expression pattern of *VcSBP* genes in different blueberry tissues. **(G–J)** Expression pattern of *VcSBP* genes in fruits at six developmental stages. Total RNAs extracted from the above different tissues and fruits at six developmental stages were used for qRT-PCR analysis. Three biological replicates and three technical replicates for each biological replicate were performed. Error bars indicate standard errors. The data were normalized against the gene *VcACTIN*.

(Hou et al., 2017). To computationally identify the SBP family members targeted by miR156/157 in blueberry, the miRNA responsive elements (MREs) were searched using the complementary sequences of six miR156/157s against the 20 *VcSBP*s and the 2 alternatively spliced variants (**Supplementary Figure 3**). It was found that 10 of the *VcSBP*s harbor one or two MRE(s) for miR156/157 (**Figure 4A**), suggesting that they have the potentials to be targeted by miR156/157. Further examination indicated that the MREs were located in the coding region of 8 *VcSBP*s (*VcSBP2*, *VcSBP6a*, *VcSBP6b*, *VcSBP8b*, *VcSBP9a*, *VcSBP9b*, *VcSBP13a*, and *VcSBP13b*) and 3′-UTR region of two *VcSBP*s (*VcSBP3* and *VcSBP5*). Noticeably, two MREs for miR156/157 were observed in the 3′-UTR region of *VcSBP3*. It is worth mentioning that, previously, *VcSBP2/SPL12* was experimentally verified to be targeted by miR156/157 (Li et al., 2020). Here, further mining of our degradome data revealed four additional miR156/157-guided cleavages of *VcSBP* transcripts, including the assembled sequences Comp14442, Comp42467, Comp25517, and Comp34004 (**Figure 4B**), which correspond to the cDNA sequences of *VcSBP9b*, *VcSBP6b*, *VcSBP5*, and *VcSBP3*, respectively. These data suggest that these *VcSBP*s might be targeted by miR156 *in vivo*.

To verify the SBP family members targeted by miR156/157s *in vivo*, genetic transformation was performed to obtain

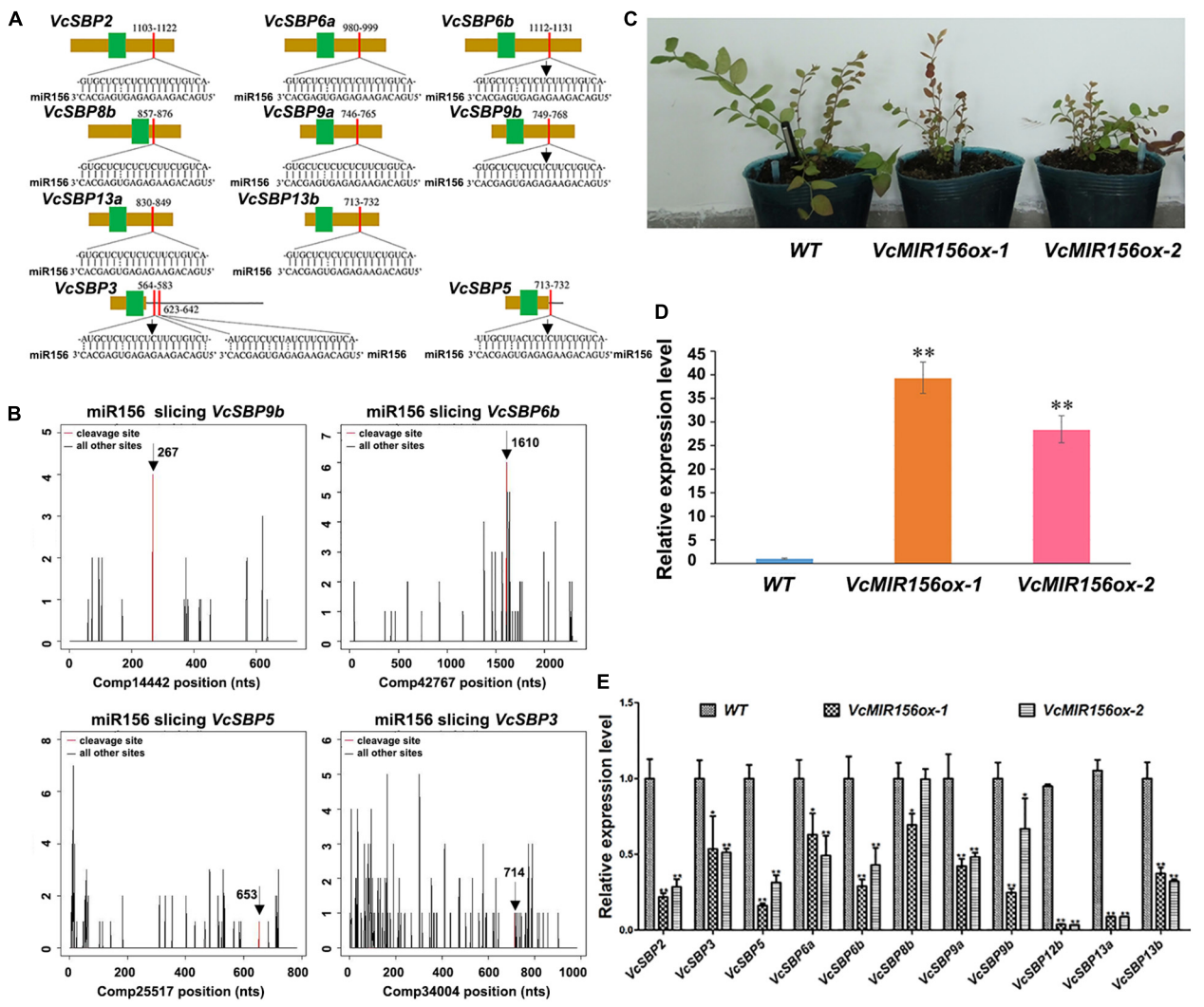

FIGURE 4 | miR156-targeted *SBP* family members in blueberry. **(A)** The diagrams of *VcSBP* sequences targeted by miR156. The brown box, black line, green box, and red box represent CDS sequence, UTR regions, SBP domains, and miR156 responsive elements, respectively. **(B)** Target plots (t-plots) of the miR156-targeted *VcSBP* transcripts (*VcSBP9b*, *VcSBP6b*, *VcSBP5*, and *VcSBP3*). The vertical lines represent the peak of different sliced *VcSBP* transcripts, and the red line denotes the peak of miR156-mediated transcript slices. The exact cleavage sites are indicated by vertical arrows in the **(A,B)**. **(C)** The photograph of transgenic blueberry lines with *VcMIR156a* overexpression and untransformed control (WT). **(D)** Expression analysis of *VcMIR156a* in transgenic blueberry lines. **(E)** The expression pattern of *VcSBP* genes in WT and transgenic blueberry lines overexpressing *VcMIR156a*. Total RNAs extracted from blueberry leaf was used for qRT-PCR analysis. Three biological replicates and three technical replicates for each biological replicate were performed. Error bars indicate standard errors. The data were normalized against the gene *VcACTIN*, and significant differences are denoted by asterisks: $*P < 0.05$, $**P < 0.01$.

transgenic blueberry lines overexpressing *VcMIR156a* (**Figure 4C**). qPCR analysis indicated that the expression of *VcMIR156a* was indeed increased in the two transgenic blueberry lines (**Figure 4D**). Furthermore, the expressions of the above 10 *SBP*s and *VcSBP12b* were examined in the transgenic blueberry lines and untransformed control. As shown in **Figure 4E**, all the examined *VcSBP* genes were significantly repressed by the *VcMIR156a* overexpression, especially *VcSBP2*, *VcSBP5*, *VcSBP6b*, *VcSBP9a*, *VcSBP12b*, *VcSBP13a*, and *VcSBP13b* with more than two-fold decreases, suggesting that these eleven *SBP* genes can be regulated through miR156/157-mediated mRNA cleavage *in vivo*.

VcSBP Family Plays Diverse Roles in Arabidopsis and Affects Chlorophyll Accumulation

To investigate the functional roles of *VcSBP* genes, transgenic Arabidopsis lines were generated for the *VcSBP* genes. Phenotypic analysis indicated that the *VcSBP* family genes performs diverse functions in Arabidopsis, mainly involved in four aspects of biological or developmental processes: flowering, leaf development, trichome formation, and chlorophyll accumulation. Overexpression of seven *VcSBP*s (*VcSBP7a/7b*, *VcSBP14a/14b*, *VcSBP3*, *VcSBP5*, and *VcSBP13a*) led to early flowering (**Figure 5A** and **Supplementary Figures 4A,B**), whereas *VcSBP8b* repressed plant flowering and trichome formation in Arabidopsis (**Figure 5A** and **Supplementary Figures 4C,D**). It was also observed that curling leaf could be arisen from overexpression of each of the three *VcSBP* genes, *VcSBP10*, *VcSBP13a*, or *VcSBP13b* (**Figure 5A** and **Supplementary Figure 4E**), while the transgenic lines overexpressing *VcSBP13a* or *VcSBP8b* showed narrow leaf (**Figures 5A,B**). Additionally, serrated leaf was observed in the transgenic lines overexpressing *VcSBP12b* or *VcSBP13a* (**Figure 5A** and **Supplementary Figure 4F**). Clearly, the *VcSBP*s in the same phylogenetic clade cannot always generate similar morphological characters (**Figure 5A**).

Previously, we reported that overexpression of *VcSBP2/SPL12* enhanced chlorophyll accumulation in Arabidopsis (Li et al., 2020). Here, we noticed that transgenic plants overexpressing each of the 4 *VcSBP* genes (*VcSBP2*, *VcSBP12b*, *VcSBP13a*, and *VcSBP8b*) clearly showed dark green leaves (**Figure 5B**). Also, a little succulence was observed for the leaves of the *VcSBP8b*-overexpressing transgenic lines (**Supplementary Figures 4C,D**). Chlorophyll contents were then determined in the transgenic Arabidopsis lines. As shown in **Figure 5C**, total chlorophyll contents in the transgenic lines overexpressing *VcSBP2*, *VcSBP12b*, or *VcSBP13a* were 1.28, 1.31, and 1.24 times higher, respectively, than that in WT. Consistently, both chlorophyll a and b were increased by 1.15–1.43 and 1.24–1.29-folds as compared to WT, respectively (**Figure 5C**). However, chlorophyll content was decreased in the *VcSBP8b*-overexpressing transgenic lines, which might be due to the succulent leaves. Furthermore, the expressions of eight chlorophyll-associated genes were examined in the *VcSBP13a*-overexpressing transgenic lines,

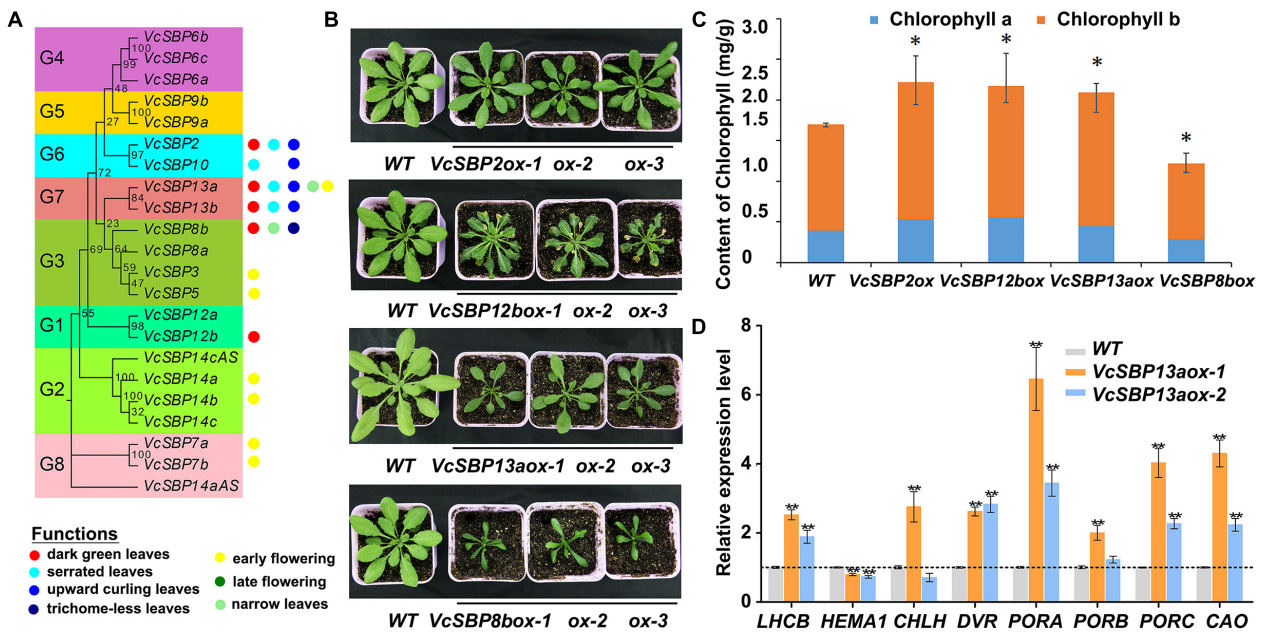

FIGURE 5 | VcSBP family play diverse roles in Arabidopsis and affects chlorophyll accumulation. **(A)** Summery of functional roles of the *VcSBP* family genes in Arabidopsis. **(B)** Photographs of transgenic Arabidopsis lines overexpressing *VcSBP* genes (*VcSBP2*, *VcSBP12b*, *VcSBP13a*, and *VcSBP8b*). **(C)** Chlorophyll contents of wild type and transgenic lines overexpressing *VcSBP2*, *VcSBP12b*, *VcSBP13a*, or *VcSBP8b*. **(D)** Expression patterns of 8 chlorophyll-associated genes (*AtHEMA1*, *AtDVR*, *AtPORA*, *AtPORB*, *AtPORC*, *AtCAO*, *AtCHLH*, and *AtLHCB*) in wild type and transgenic lines overexpressing *VcSBP13a*. Total RNAs were exacted from 14-day-old transgenic seedlings and wild type. Values were normalized against the gene *AtACTIN8*. Error bars in **(C,D)** indicate standard errors of three biological and technical replicates, and significant differences are denoted by asterisks: $^{*}P < 0.05$, $^{**}P < 0.01$.

including 7 chlorophyll biosynthetic genes (*AtHEMA1*, *AtDVR*, *AtPORA*, *AtPORB*, *AtPORC*, *AtCAO*, and *AtCHLH*) and one chlorophyll-binding protein gene (*AtLHCB*). Consequently, all the genes were significantly upregulated by overexpression of *VcSBP13a* except *AtHEMA1* that showed a slight decrease (**Figure 5D**). Especially, *AtPORA* was remarkably increased by 10-folds (**Figure 5D**). These results indicated that *VcSBP*s affect chlorophyll accumulation via regulating the expression of chlorophyll-associated genes in Arabidopsis. Since *SBP* family is transcriptionally regulated by miR156, the expressions of the above eight chlorophyll biosynthetic genes were examined in the *VcMIR156a*-overexpressing transgenic Arabidopsis. Consistently, *AtDVR* and *AtPORC* were significantly repressed by *VcMIR156a* overexpression (**Supplementary Figure 5**).

SBP Family Might Affect the Expression of *VcLHCB1* via Targeting Its Promoter in Blueberry

Increasing evidence indicated that SBPs are able to bind to the consensus sequence TNCGTACAA with GTAC as its essential core (Kropat et al., 2005). To provide some clues for understanding the targets of SBPs in blueberry, the potential genome-wide binding sites were searched using the consensus sequence against the *V. corymbosum* cv. Draper v1.0 genome. Consequently, 2568 genes were found to harbor the potential binding motif of SBP proteins in their promoter regions (**Supplementary Table 4**), suggesting that they are possible targets of SBP proteins in blueberry. The potential targets were classified into five groups based on their functional roles, including transcription, DNA-or-RNA-related; metabolism defense or protein binding; cellular process; synthesis, catalysis or modification; biological process unknown (**Figure 6A**). Further examination indicated that the distribution of the potential binding sites in the promoter regions varied among the five groups. The density of the binding sites over the target genes in the groups G5, G3, and G2 peaked around ∼900, 1100, and 1200 bp upstream of their TSSs, while no obvious peak was found for the target genes in the groups G4 and G1 (**Figure 6B**).

Consistent with the above results that *VcSBP*s affect chlorophyll accumulation in Arabidopsis, nine chlorophyll-associated genes were found to harbor the potential binding site of SBP proteins, including three *LIGHT HARVESTING CHLOROPHYLL A/B BINDING PROTEIN*s (*VcLHCB*s), two *LOW QUANTUM YIELD OF PHOTOSYSTEM II1* (*VcLQY1*), 3 *CHAPERONE-LIKE PROTEIN OF POR1-like* (*VcCPP1*), and one *LOW PSII ACCUMULATION 3* (*VcLPA3*). Subsequently, five genes were chosen as representatives to examine their expression patterns in transgenic blueberry plants overexpressing *VcMIR156a* where 11 *VcSBP*s were transcriptionally repressed (**Figure 4E**). As shown in **Figure 6C**, the expressions of *VcLPA3* and two *VcLHCB*s were significantly downregulated by *MIR156a* overexpression in blueberry, whereas no significant change was observed for *VcLQY1* and *VcCPP1*. Furthermore, the expression patterns of their corresponding Arabidopsis homologs (such as *AtLHCB2.1*, *AtLHCB2.2*, *AtLHCB5*, *AtLQY1*, *AtCPP1*, and *AtLPA3*) were investigated in transgenic Arabidopsis with *VcSBP13a* overexpression. Consequently, all these homologous genes were significantly promoted by *VcSBP13a* overexpression (**Figure 6D**).

To examine if the SBP proteins bind to these genes, VcSBP13a and two *VcLHCB*s (*VcLHCB1* and *VcLHCB2*) were separately chosen as representatives of baits and preys to perform Y1H analysis. Sequence analysis indicated that the promoter region of *VcLHCB1* contained two typical binding sites (TNCGTACAA element), whereas only GTAC elements were observed in the promoter region of *VcLHCB2* (**Figure 6E**). As shown in **Figure 6F**, like the positive control (*pB42AD-AtRVE8/placZi-AtPRR5*), strong blue colonies were observed when VcSBP13a acts as bait and the fragment of *VcLHCB1* promoter as prey. In contrast, very light blue appeared in the colonies containing *pB42AD* as bait and *pLacZi-pVcLHCB1* as prey, and no blue color was shown for the other two negative controls (*pB42AD/pLacZi* and *pB42AD-VcSBP13a/pLacZi*). These results indicated that physical interaction occurred between VcSBP13a and the *VcLHCB1* promoter. However, no blue color was observed in the colonies when *pB42AD-VcSBP13a* acted as bait and *pLacZi-pVcLHCB2* as prey.

DISCUSSION

SBP genes belong to a small family of plant-specific transcription factors. In the present study, 20 *SBP* genes were identified in blueberry, and the number of VcSBP family members is similar to the ones in Petunia (21), Tartary buckwheat (24), grape (17), and pear (19) (Hou et al., 2013; Li et al., 2013; Zhou et al., 2018; Liu et al., 2019), supporting the notion that the number of *SBP* genes in different plant species is relatively stable during evolution (Liu et al., 2019). Gene family generally arises from gene duplication during evolution, therefore leading to the acquisition of neofunctionalizations and subfunctionalizations as well as the emergence of backup or redundant genes (Preston and Hileman, 2013; Wang and Wang, 2015). Among the 20 identified *VcSBP*s, only two gene pairs (*VcSPB7a* and *VcSBP7b*, *VcSBP14b*, and *VcSBP14c*) might have been derived from segmental duplication, and one pair (*VcSBP8a* and *VcSBP8b*) from tandem amplification (**Figure 1C**). Noticeably, the *VcSBP* genes in the same group in the phylogenetic tree showed relatively low identity (8–48%) except the two segmental duplication pairs (*VcSPB7a* and *VcSBP7b*, *VcSBP14b*, and *VcSBP14c*, **Figures 1C**, **2A**), suggesting that most *VcSBP*s might be single-copy genes with functional specificity. However, it was estimated that at least three rounds of whole-genome duplication occurred during the evolution of blueberry species (Wang et al., 2020), which are supposed to facilitate the generation of multiple copy genes. It can be explained by at least two reasons: (1) it is still possible that the number of *VcSBP* genes might have been underestimated since the identification of *SBP* genes was conducted on the basis of available transcriptome data; (2) *VcSBP*s might belong to the duplication-resistant genes, which generally return to single-copy status through the duplication-resistant system or genetic drift after suffering duplication events (Wang et al., 2020).

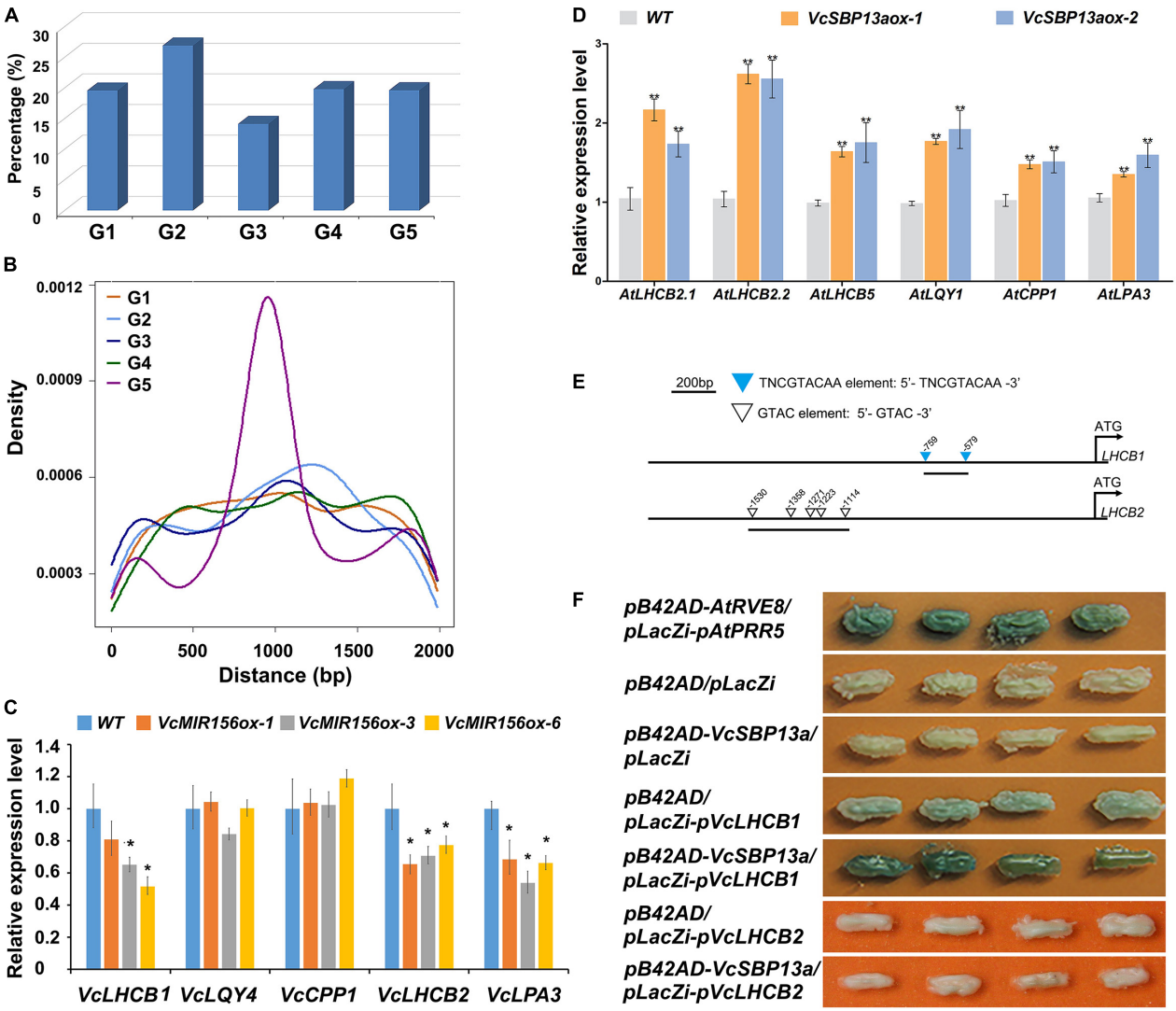

FIGURE 6 | Computational and experimental analysis of the target client(s) of SBP(s) in blueberry. **(A)** The categories of putative target clients. Five categories were classified in terms of their functional roles, including transcription, DNA-or-RNA-related (G1); metabolism defense or protein binding (G2); cellular process (G3); synthesis, catalysis, or modification (G4); and biological process unknown (G5). **(B)** The density of SBP binding sites in the promoter regions of each group of target clients. **(C)** The expression patterns of chlorophyll-associated genes (*VcLHCB1*, *VcLQY4*, *VcCPP1*, *VcLHCB2*, and *VcLPA3*) in WT and transgenic blueberry lines overexpressing *VcMIR156a*. **(D)** The expression patterns of chlorophyll-associated genes (*AtLHCB2.1*, *AtLHCB2.2*, *AtLHCB5*, *AtLQY4*, *AtCPP1*, and *AtLPA3*) in WT and transgenic Arabidopsis lines overexpressing *VcSBP13a*. Values were normalized against the gene *VcACTIN* **(C)** or *AtACTIN8* **(D)**. Error bars indicate standard errors of three biological and technical replicates, and significant differences are denoted by asterisks: $^{*}P < 0.05$, $^{**}P < 0.01$. **(E)** The schematic diagrams of the promoters of *VcLHCB1* and *VcLHCB2*. The downward triangles indicate the elements of TNCGTACAA and GTAC, and the number shows the start site of each element. **(F)** Interaction assay between *VcSBP13a* and the promoters of *VcLHCB1* and *VcLHCB2* using the yeast one-hybrid assay. *pB42AD/pLacZi-pVcLHCB1*, *pB42AD/pLacZi-pVcLHCB2*, *pB42AD/pLacZi*, and *pB42AD-VcSBP13a/pLacZi* were set as negative controls; and *pB42AD-AtRVE8/placZi-AtPRR5* as positive control.

SBP family performs diverse functions during plant growth and development. Here, we presented four aspects of evidence to show the functional diversity of the *SBP* family in blueberry. Firstly, it has been proposed that the diversification of gene structures and conservation of motifs may be tightly associated with the functional evolution of *SBP* genes (Salinas et al., 2012; Hou et al., 2013; Li et al., 2013; Li and Lu, 2014; Shalom et al., 2015; Liu et al., 2019). In the present study, it was revealed that the majority of VcSBP family members belonging to the same phylogenetic group showed similar motif compositions and gene structures, while the diversity in motif compositions and gene structures was observed between the SBP family members in different phylogenetic groups (**Figure 2**). These observations support a scenario that the SBP family underwent functional conservation and diversification during evolution (Preston and Hileman, 2013; Wang and Wang, 2015; Zhang et al., 2015). Secondly, the spatio-temporal expression is generally thought as key contributors to functional specificity for a gene family

(Paul et al., 2012; Hasan et al., 2017). *VcSBP* genes displayed tissue-specific and fruit development stage-specific expression patterns (**Figure 3**), implying that the *VcSBP* family might exert diverse functions in blueberry. Thirdly, the SBP family can be regulated via miR156-guided transcript cleavage or translational repression, and a subset of *SBP*s have been proved to be targeted by miR156 through recognizing MREs on their transcripts, for example, 11 out of the 17 *SBP*s in Arabidopsis, 15 out of the 27 *SBP*s in apple and 7 out of the 19 *SBP*s in pear (Li et al., 2013; Zhang et al., 2015; Qian et al., 2017). In the present study, several members of the *VcSBP* family were computationally and experimentally demonstrated to be targets of miR156 (**Figure 4**), implying that a subset of *VcSBP*s are able to form a regulatory hub with miR156, therefore exerting vital functions during blueberry growth and development. Lastly, overexpression of *VcSBP*s in Arabidopsis gave rise to multiple morphological phenotypes (**Figure 5** and **Supplementary Figure 4**). Thus, our results provided an overall framework for understanding the functional diversity of *VcSBP* genes, which will contribute to the genetic improvement of the agronomic traits of blueberry.

The SBP family acts as pivotal regulators of diverse biological and physiological processes in plants. In the present study, functional analysis in Arabidopsis indicated that *VcSBP* genes might be involved in multiple developmental processes such as leaf shape regulation (serrated leaf formation, *VcSBP12b/13a*; narrow leaf, *VcSBP8b/13a*), trichome formation (*VcSBP8b*), and flowering time control (*VcSBP7a/7b/14a/14b/3/5/13a*, **Figure 5** and **Supplementary Figure 4**). These observations are consistent with previous reports in Arabidopsis. For example, *AtSPL10* overexpression causes narrow leaf in Arabidopsis (Gao et al., 2018); *AtSPL3/4/5* exert important functions in regulating Arabidopsis flowering and developmental transition (Jung et al., 2016; Xu et al., 2016); *AtSPL3/4/5/8/9/10/13* affect trichome formation (Yu et al., 2010); and loss-of-function mutation of *AtSPL14* increases the number of leaf hydathodes and enhances leaf margin serration (Stone et al., 2005). Thus, our results support the notion that the functionality of the SBP family proteins is highly conserved among distinct plant species (Preston and Hileman, 2013). However, not all the *VcSBP*s display the same functional roles as their counterparts in other plant species. For instance, overexpression of *VcSBP8b* leads to a very narrow leaf in Arabidopsis (**Figure 5B**), whereas its Arabidopsis counterpart *AtSPL8* fails to generate similar leaf morphology, and it is *AtSPL10* instead that was reported to modulate leaf morphology (Gao et al., 2018). Previous studies also indicated that mutation of *LG1*, the closest homolog of Arabidopsis *AtSPL8*, in maize, rice, and barley gave rise to the lack of ligules and auricles (Lee et al., 2007; Wang and Wang, 2015), whereas in Arabidopsis mutation of *AtSPL8* fails to cause a similar structure of ligules. Thus, it seems that it is not always possible to foretell the functional roles of individual *SBP* genes based on homology, although the *SBP* family as a whole shows functional conservation across diverse plant species. More interestingly, overexpression of three *VcSBP* genes (*VcSBP10/13a/13b*) in Arabidopsis causes the formation of curling leaves (**Supplementary Figure 4E**). Previous report indicated that mutation in *rSPL13* led to an up-curled leaf phenotype in alfalfa (Gao et al., 2018). Nevertheless, no evidence shows the formation of curling leaves by being members of the SBP family in Arabidopsis. Thus, it appears that *SBP* family might show species-dependent functions or novel function in some specific plant species.

Several studies indicated that the SBP family plays important roles during fruit ripening. For example, *VmTDR4* (a SQUAMOSA-class *MADS-box* gene) is positively involved in the regulation of anthocyanin accumulation during bilberry fruit ripening (Jaakola et al., 2010), while *MaSPL16* in banana regulates carotenoid biosynthesis through promoting the expression of *MaLCYB*s (Zhu et al., 2020). Likewise, *SlSPL-CNR*, an SBP transcription factor in tomato, is mainly expressed in ripening fruits and serves as a positive player in the regulation of fruit ripening and cell death (Lai et al., 2020). In the present study, five *VcSBP* genes were found to be expressed at relatively low levels at three early stages of fruit development and significantly increased during fruit ripening (**Figure 3I**), suggesting that they might act as regulatory hubs to control fruit ripening in blueberry. In contrast, 10 *VcSBP* genes were highly expressed at three early stages of fruit development and dramatically decreased to a low level when fruit initiates ripening (**Figure 3G,H**), which is consistent with previous reports that the expressions of *VvSPL6/10/13* were gradually decreased as grape berry develops and ripens (Cui et al., 2018), while the *FvSPL*s were transcriptionally decreased during strawberry fruit ripening (Xiong et al., 2018). These results suggest that the 10 *VcSBP*s might be required for fruit development and suppressed during fruit ripening in blueberry.

Generally, the development and ripening of fleshy fruits are accompanied by a wide range of changes at cellular, molecular and metabolic levels, including fruit enlargement, degreening, accumulation of pigments, softening, etc. Previously, we revealed that *VcSBP2/SPL12* affects the accumulation of chlorophylls in Arabidopsis (Li et al., 2020). In the present study, the contents of chlorophyll a and b were found to be increased by the overexpression of at least three *SBP* genes (*VcSBP2*, *VcSBP12a*, and *VcSBP13a*) in Arabidopsis (**Figure 5**). Moreover, the chlorophyll biosynthetic genes in Arabidopsis were indeed elevated by *VcSBP13a* overexpression (**Figure 5D**). These observations indicated that a subset of *VcSBP*s might be involved in the regulation of chlorophyll accumulation. Previous studies have revealed that SBP family proteins can directly interact with their clients (for example, *AtFUL*, *AtSOC1*, *AtDFR*, *AtAP1*, *MdWRKY100*, *MaLCYB1.1*, and *MaLCYB1.2*, *MADS5,* and *MADS32*), thereby regulating diverse biological processes in plants such as flowering, inflorescence formation, biosynthesis of secondary metabolites, root regeneration, and response to stress (Yamaguchi et al., 2009; Gou et al., 2019; Ma et al., 2020; Zhu et al., 2020). However, the targets of SBP proteins associated with chlorophyll accumulation have remained to be found. Our Y1H assay showed that VcSBP13a could physically bind to the promoter region of an *LHCB* gene in blueberry (**Figure 6**). Thus, we proposed that VcSBPs are able to positively regulate the expressions of chlorophyll-associated genes (at least *VcLHCB1*) directly through binding to their promoter regions to affect chlorophyll accumulation in blueberry.

In conclusion, the SBP family was systematically identified and functionally characterized in blueberry, and they show conservation and divergence in characteristics and functional roles across plant species. Based on the targets and functional roles of *VcSBP*s as well as their expression patterns, we propose that a subset of *VcSBP*s might be involved in the regulation of chlorophyll accumulation directly through targeting to the chlorophyll-associated genes such as *VcLHCB1*. These findings provide the first comprehensive understandings of the features and functional diversity of the SBP family in blueberry, which will facilitate their utilization in the improvement of the agronomic traits of blueberry.

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

SB and XL designed the experiments. XX, SY, BS, HL, and PY performed the experiments. JW and SL performed the data analyzes. SB, XL, and YC wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China [Grant number 31872075, 2019–2022].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.703994/full#supplementary-material

Supplementary Figure 1 | Phylogenetic analysis of *SBP* genes in blueberry and Arabidopsis. The CDS sequences of *AtSBP*s were downloaded from the TAIR website (www.arabidopsis.org). A neighbor-joining tree was generated with the MEGA7 software using the CDS sequences of the *SBP* genes in blueberry and Arabidopsis.

Supplementary Figure 2 | Phylogenetic relationship of VcSBP proteins with the SBPs in other plant species. A maximum likelihood tree was generated with the MEGA X software using the putative amino acid sequences of 101 SBP proteins, which were clustered into eight groups (G1-G6). The SBP proteins in the same species are represented with the same symbol: blue check, *V. corymbosum*; orange circle, *Vitis vinifera*; yellow square, *Malus domestica*; pink star, *Solanum lycopersicum*; gray triangle, *Arabidopsis thaliana*.

Supplementary Figure 3 | Alignment of miR156 complementary sequences within *VcSBP* genes. **(A)** Complementary sequences are within coding regions and the 3′UTR. Reverse complement sequences of the mature vco-miR156a, vco-miR156b-5p, vco-miR156c-5p, vco-miR156e-5p, vco-miR156g-5p, and vco-miR157a genes are shown below the alignment for comparison. **(B)** The sequence logo of miR156 responsive elements in *VcSBP* genes. The total height of each stack represents the conservation degree of each position, while the height of the letters within each stack indicates the relative frequency of the corresponding amino acid.

Supplementary Figure 4 | Phenotype of the transgenic Arabidopsis individually overexpressing the *VcSBP* family genes. Early **(A, B)** – and – late **(C)** flowering phenotypes of the transgenic Arabidopsis overexpressing the gene *VcSBP7a*, *VcSBP7b*, *VcSBP14a*, *VcSBP14b*, *VcSBP3*, *VcSBP5*, *VcSBP13a*, or *VcSBP8b*. **(D)** Less trichome phenotype of the transgenic Arabidopsis overexpressing the gene *VcSBP7a*. **(E)** Curled leaf phenotype of the transgenic Arabidopsis overexpressing the gene *VcSBP10*, *VcSBP13a*, or *VcSBP13b*. **(F)** Serrated leaf phenotype of the transgenic Arabidopsis overexpressing the gene *VcSBP13a* or *VcSBP12b*.

Supplementary Figure 5 | Expression patterns of 8 chlorophyll-associated genes (*AtLHCB*, *AtHEMA1*, *AtCHLH*, *AtDVR*, *AtPORA*, *AtPORB*, *AtPORC*, and *AtCAO*) in wild type and transgenic lines overexpressing *VcMIR156a*. Total RNAs were exacted from 7-day-old transgenic seedlings and wild type. Values were normalized against the gene *AtACTIN8*. Error bars indicate standard errors of three biological and technical replicates, and significant differences are denoted by asterisks: $^{*}P < 0.05$, $^{**}P < 0.01$.

Supplementary Table 1 | Primers used in the study.

Supplementary Table 2 | Characterization of the SBP family in blueberry.

Supplementary Table 3 | Similarity and identity between different VcSBP family members.

Supplementary Table 4 | The predicted targets of the VcSBP family proteins in blueberry.

REFERENCES

Barrera-Rojas, C. H., Rocha, G. H. B., Polverari, L., Brito, D. A. P., Batista, D. S., Notini, M. M., et al. (2020). miR156-targeted SPL10 controls Arabidopsis root meristem activity and root-derived de novo shoot regeneration via cytokinin responses. *J. Exp. Bot.* 71, 934–950. doi: 10.1093/jxb/erz475

Bhogale, S., Mahajan, A. S., Natarajan, B., Rajabhoj, M., Thulasiram, H. V., and Banerjee, A. K. (2014). MicroRNA156: a potential graft-transmissible MicroRNA that modulates plant architecture and tuberization in Solanum tuberosum ssp andigena. *Plant Physiol.* 164, 1011–1027. doi: 10.1104/pp.113.230714

Birkenbihl, R. P., Jach, G., Saedler, H., and Huijser, P. (2005). Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. *J. Mol. Biol.* 352, 585–596. doi: 10.1016/j.jmb.2005.07.013

Colle, M., Leisner, C. P., Wai, C. M., Ou, S., Bird, K. A., Wang, J., et al. (2019). Haplotype-phased genome and evolution of phytonutrient pathways of tetraploid blueberry. *Gigascience* 8:giz012. doi: 10.1093/gigascience/giz012

Cui, M., Wang, C., Zhang, W., Pervaiz, T., Haider, M. S., Tang, W., et al. (2018). Characterization of Vv-miR156: Vv-SPL pairs involved in the modulation of grape berry development and ripening. *Mol. Genet. Genomics* 293, 1333–1354. doi: 10.1007/s00438-018-1462-1

Gao, R., Gruber, M. Y., Amyot, L., and Hannoufa, A. (2018). SPL13 regulates shoot branching and flowering time in Medicago sativa. *Plant Mol. Biol.* 96, 119–133. doi: 10.1007/s11103-017-0683-8

Gou, J., Tang, C., Chen, N., Wang, H., Debnath, S., Sun, L., et al. (2019). SPL7 and SPL8 represent a novel flowering regulation mechanism in switchgrass. *New Phytol.* 222, 1610–1623. doi: 10.1111/nph.15712

Gupta, V., Estrada, A. D., Blakley, I., Reid, R., Patel, K., Meyer, M. D., et al. (2015). RNA-Seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. *Gigascience* 4:5. doi: 10.1186/s13742-015-0046-9

Hasan, M. A., Ahmad, S., and Molla, M. K. (2017). Protein subcellular localization prediction using multiple kernel learning based support vector machine. *Mol. Biosyst.* 13, 785–795. doi: 10.1039/c6mb00860g

Hou, H., Li, J., Gao, M., Singer, S. D., Wang, H., Mao, L., et al. (2013). Genomic organization, phylogenetic comparison and differential expression of the SBP-box family genes in grape. *PLoS One* 8:e59358. doi: 10.1371/journal.pone.0059358

Hou, Y., Zhai, L., Li, X., Xue, Y., Wang, J., Yang, P., et al. (2017). Comparative analysis of fruit ripening-related miRNAs and their targets in blueberry using small RNA and degradome sequencing. *Int. J. Mol. Sci.* 18:2767. doi: 10.3390/ijms18122767

Hou, Y. M., Li, H. X., Zhai, L. L., Xie, X., Li, X. Y., and Bian, S. M. (2020). Identification and functional characterization of the Aux/IAA gene VcIAA27 in blueberry. *Plant Signal. Behav.* 15:1700327. doi: 10.1080/15592324.2019.1700327

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

Jung, J. H., Lee, H. J., Ryu, J. Y., and Park, C. M. (2016). SPL3/4/5 integrate developmental aging and photoperiodic signals into the FT-FD module in Arabidopsis flowering. *Mol. Plant* 9, 1647–1659. doi: 10.1016/j.molp.2016.10.014

Klein, J., Saedler, H., and Huijser, P. (1996). A new family of DNA binding proteins includes putative transcriptional regulators of the Antirrhinum majus floral meristem identity gene SQUAMOSA. *Mol. Gen. Genet.* 250, 7–16. doi: 10.1007/BF02191820

Kropat, J., Tottey, S., Birkenbihl, R. P., Depege, N., Huijser, P., and Merchant, S. (2005). A regulator of nutritional copper signaling in Chlamydomonas is an SBP domain protein that recognizes the GTAC core of copper response element. *Proc. Natl. Acad. Sci. U. S. A.* 102, 18730–18735. doi: 10.1073/pnas.0507693102

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

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

Lai, T., Wang, X., Ye, B., Jin, M., Chen, W., Wang, Y., et al. (2020). Molecular and functional characterization of the SBP-box transcription factor SPL-CNR in tomato fruit ripening and cell death. *J. Exp. Bot.* 71, 2995–3011. doi: 10.1093/jxb/eraa067

Lee, J., Park, J. J., Kim, S. L., Yim, J., and An, G. (2007). Mutations in the rice liguleless gene result in a complete loss of the auricle, ligule, and laminar joint. *Plant Mol. Biol.* 65, 487–499. doi: 10.1007/s11103-007-9196-1

Lei, K. J., Lin, Y. M., Ren, J., Bai, L., Miao, Y. C., An, G. Y., et al. (2016). Modulation of the phosphate-deficient responses by MicroRNA156 and its targeted SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3 in Arabidopsis. *Plant Cell Physiol.* 57, 192–203. doi: 10.1093/pcp/pcv197

Leister, D. (2004). Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance genes. *Trends Genet.* 20, 116–122. doi: 10.1016/j.tig.2004.01.007

Li, C., and Lu, S. (2014). Molecular characterization of the SPL gene family in Populus trichocarpa. *BMC Plant Biol.* 14:131. doi: 10.1186/1471-2229-14-131

Li, J., Hou, H., Li, X., Xiang, J., Yin, X., Gao, H., et al. (2013). Genome-wide identification and analysis of the SBP-box family genes in apple (Malus x domestica Borkh.). *Plant Physiol. Biochem.* 70, 100–114. doi: 10.1016/j.plaphy.2013.05.021

Li, L., Zhang, H., Liu, Z., Cui, X., Zhang, T., Li, Y., et al. (2016). Comparative transcriptome sequencing and de novo analysis of Vaccinium corymbosum during fruit and color development. *BMC Plant Biol.* 16:223. doi: 10.1186/s12870-016-0866-5

Li, X., Hou, Y., Xie, X., Li, H., Li, X., Zhu, Y., et al. (2020). A blueberry MIR156a-SPL12 module coordinates the accumulation of chlorophylls and anthocyanins during fruit ripening. *J. Exp. Bot.* 71, 5976–5989. doi: 10.1093/jxb/eraa327

Liu, M., Sun, W., Ma, Z., Huang, L., Wu, Q., Tang, Z., et al. (2019). Genome-wide identification of the SPL gene family in Tartary Buckwheat (Fagopyrum tataricum) and expression analysis during fruit development stages. *BMC Plant Biol.* 19:299. doi: 10.1186/s12870-019-1916-6

Luo, L., Li, W., Miura, K., Ashikari, M., and Kyozuka, J. (2012). Control of tiller growth of rice by OsSPL14 and Strigolactones, which work in two independent pathways. *Plant Cell Physiol.* 53, 1793–1801. doi: 10.1093/pcp/pcs122

Ma, Y., Xue, H., Zhang, F., Jiang, Q., Yang, S., Yue, P., et al. (2020). The miR156/SPL module regulates apple salt stress tolerance by activating MdWRKY100 expression. *Plant Biotechnol. J.* 19, 311–323. doi: 10.1111/pbi.13464

Manning, K., Tor, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., et al. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* 38, 948–952. doi: 10.1038/ng1841

Nguyen, S. T., Greaves, T., and Mccurdy, D. W. (2017). Heteroblastic development of transfer cells is controlled by the microRNA miR156/SPL module. *Plant Physiol.* 173, 1676–1691. doi: 10.1104/pp.16.01741

Paul, A. L., Denison, F. C., Schultz, E. R., Zupanska, A. K., and Ferl, R. J. (2012). 14-3-3 phosphoprotein interaction networks – does isoform diversity present functional interaction specification? *Front. Plant Sci.* 3:190. doi: 10.3389/fpls.2012.00190

Plunkett, B. J., Espley, R. V., Dare, A. P., Warren, B. A. W., Grierson, E. R. P., Cordiner, S., et al. (2018). MYBA from blueberry (Vaccinium Section Cyanococcus) is a subgroup 6 type R2R3MYB transcription factor that activates anthocyanin production. *Front. Plant Sci.* 9:1300. doi: 10.3389/fpls.2018.01300

Preston, J. C., and Hileman, L. C. (2013). Functional evolution in the plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) gene family. *Front. Plant Sci.* 4:80. doi: 10.3389/fpls.2013.00080

Preston, J. C., Jorgensen, S. A., Orozco, R., and Hileman, L. C. (2016). Paralogous SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) genes differentially regulate leaf initiation and reproductive phase change in petunia. *Planta* 243, 429–440. doi: 10.1007/s00425-015-2413-2

Qi, X., Ogden, E. L., Ehlenfeldt, M. K., and Rowland, L. J. (2019). Dataset of de novo assembly and functional annotation of the transcriptome of blueberry (Vaccinium spp.). *Data Brief.* 25:104390. doi: 10.1016/j.dib.2019.104390

Qian, M., Ni, J., Niu, Q., Bai, S., Bao, L., Li, J., et al. (2017). Response of miR156-SPL module during the red peel coloration of bagging-treated chinese sand pear (Pyrus pyrifolia Nakai). *Front. Physiol.* 8:550. doi: 10.3389/fphys.2017.00550

Routray, W., and Orsat, V. (2011). Blueberries and their anthocyanins: factors affecting biosynthesis and properties. *Compr. Rev. Food Sci. F.* 10, 303–320. doi: 10.1111/j.1541-4337.2011.00164.x

Rowland, L. J., Alkharouf, N., Darwish, O., Ogden, E. L., Polashock, J. J., Bassil, N. V., et al. (2012). Generation and analysis of blueberry transcriptome sequences from leaves, developing fruit, and flower buds from cold acclimation through deacclimation. *BMC Plant Biol.* 12:46. doi: 10.1186/1471-2229-12-46

Salinas, M., Xing, S., Hohmann, S., Berndtgen, R., and Huijser, P. (2012). Genomic organization, phylogenetic comparison and differential expression of the SBP-box family of transcription factors in tomato. *Planta* 235, 1171–1184. doi: 10.1007/s00425-011-1565-y

Schwarz, S., Grande, A. V., Bujdoso, N., Saedler, H., and Huijser, P. (2008). The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in Arabidopsis. *Plant Mol. Biol.* 67, 183–195. doi: 10.1007/s11103-008-9310-z

Shalom, L., Shlizerman, L., Zur, N., Doron-Faigenboim, A., Blumwald, E., and Sadka, A. (2015). Molecular characterization of SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) gene family from Citrus and the effect of fruit load on their expression. *Front. Plant Sci.* 6:389. doi: 10.3389/fpls.2015.00389

Shikata, M., Koyama, T., Mitsuda, N., and Ohme-Takagi, M. (2009). Arabidopsis SBP-box genes SPL10, SPL11 and SPL2 control morphological change in association with shoot maturation in the reproductive phase. *Plant Cell Physiol.* 50, 2133–2145. doi: 10.1093/pcp/pcp148

Silva, G. F. F., Silva, E. M., M da Azevedo, S., Guivin, M. A., Ramiro, D. A., et al. (2014). microRNA156-targeted SPL/SBP box transcription factors regulate tomato ovary and fruit development. *Plant J.* 78, 604–618. doi: 10.1111/tpj.12493

Song, G. Q., and Gao, X. (2017). Transcriptomic changes reveal gene networks responding to the overexpression of a blueberry DWARF AND DELAYED FLOWERING 1 gene in transgenic blueberry plants. *BMC Plant Biol.* 17:106. doi: 10.1186/s12870-017-1053-z

Song, G. Q., and Sink, K. C. (2006). Blueberry (Vaccinium corymbosum L.). *Methods Mol. Biol.* 344, 263–272. doi: 10.1385/1-59745-131-2:263

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

Stone, J. M., Liang, X., Nekl, E. R., and Stiers, J. J. (2005). Arabidopsis AtSPL14, a plant-specific SBP-domain transcription factor, participates in plant development and sensitivity to fumonisin B1. *Plant J.* 41, 744–754. doi: 10.1111/j.1365-313X.2005.02334.x

Usami, T., Horiguchi, G., Yano, S., and Tsukaya, H. (2009). The more and smaller cells mutants of Arabidopsis thaliana identify novel roles for SQUAMOSA PROMOTER BINDING PROTEIN-LIKE genes in the control of heteroblasty. *Development* 136, 955–964. doi: 10.1242/dev.028613

Walworth, A. E., Chai, B., and Song, G. Q. (2016). Transcript profile of flowering regulatory genes in VcFT-overexpressing blueberry plants. *PLoS One* 11:e0156993. doi: 10.1371/journal.pone.0156993

Wang, H., and Wang, H. Y. (2015). The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Mol. Plant* 8, 677–688. doi: 10.1016/j.molp.2015.01.008

Wang, S., Wu, K., Qian, Q., Liu, Q., Li, Q., Pan, Y., et al. (2017). Non-canonical regulation of SPL transcription factors by a human OTUB1-like deubiquitinase defines a new plant type rice associated with higher grain yield. *Cell Res.* 27, 1142–1156. doi: 10.1038/cr.2017.98

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., et al. (2012). Control of grain size, shape and quality by OsSPL16 in rice. *Nat. Genet.* 44, 950–954. doi: 10.1038/ng.2327

Wang, Y., Nie, F., Shahid, M. Q., and Baloch, F. S. (2020). Molecular footprints of selection effects and whole genome duplication (WGD) events in three blueberry species: detected by transcriptome dataset. *BMC Plant Biol.* 20:250. doi: 10.1186/s12870-020-02461-w

Xie, Y., Zhou, Q., Zhao, Y., Li, Q., Liu, Y., Ma, M., et al. (2020). FHY3 and FAR1 integrate light signals with the miR156-SPL module-mediated aging pathway to regulate arabidopsis flowering. *Mol. Plant* 13, 483–498. doi: 10.1016/j.molp.2020.01.013

Xie, Z., Su, Z., Wang, W., Guan, L., Bai, Y., Zhu, X., et al. (2019). Characterization of VvSPL18 and its expression in response to exogenous hormones during grape berry development and ripening. *Cytogenet. Genome Res.* 159, 97–108. doi: 10.1159/000503912

Xing, S., Salinas, M., Hohmann, S., Berndtgen, R., and Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in Arabidopsis. *Plant Cell* 22, 3935–3950. doi: 10.1105/tpc.110.079343

Xiong, J. S., Zheng, D., Zhu, H. Y., Chen, J. Q., Na, R., and Cheng, Z. M. (2018). Genome-wide identification and expression analysis of the SPL gene family in woodland strawberry Fragaria vesca. *Genome* 61, 675–683. doi: 10.1139/gen-2018-0014

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

Yamaguchi, A., Wu, M. F., Yang, L., Wu, G., Poethig, R. S., and Wagner, D. (2009). The MicroRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1. *Dev. Cell* 17, 268–278. doi: 10.1016/j.devcel.2009.06.007

Ye, B. B., Shang, G. D., Pan, Y., Xu, Z. G., Zhou, C. M., Mao, Y. B., et al. (2020). AP2/ERF transcription factors integrate age and wound signals for root regeneration. *Plant Cell* 32, 226–241. doi: 10.1105/tpc.19.00378

Yu, N., Cai, W. J., Wang, S., Shan, C. M., Wang, L. J., and Chen, X. Y. (2010). Temporal control of trichome distribution by microRNA156-targeted SPL genes in Arabidopsis thaliana. *Plant Cell* 22, 2322–2335. doi: 10.1105/tpc.109.072579

Yu, N., Niu, Q. W., Ng, K. H., and Chua, N. H. (2015). The role of miR156/SPLs modules in Arabidopsis lateral root development. *Plant J.* 83, 673–685. doi: 10.1111/tpj.12919

Zhang, S. D., Ling, L. Z., and Yi, T. S. (2015). Evolution and divergence of SBP-box genes in land plants. *BMC Genomics* 16:787. doi: 10.1186/s12864-015-1998-y

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

Zhou, Q., Zhang, S., Chen, F., Liu, B., Wu, L., Li, F., et al. (2018). Genome-wide identification and characterization of the SBP-box gene family in Petunia. *BMC Genomics* 19:193. doi: 10.1186/s12864-018-4537-9

Zhu, L. S., Liang, S. M., Chen, L. L., Wu, C. J., Wei, W., Shan, W., et al. (2020). Banana MaSPL16 modulates carotenoid biosynthesis during fruit ripening through activating the transcription of lycopene beta-cyclase genes. *J. Agr. Food Chem.* 68, 1286–1296. doi: 10.1021/acs.jafc.9b07134

Zifkin, M., Jin, A., Ozga, J. A., Zaharia, L. I., Schernthaner, J. P., Gesell, A., et al. (2012). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. *Plant Physiol.* 158, 200–224. doi: 10.1104/pp.111.180950

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 in Plant Science

MINI REVIEW
published: 05 July 2021
doi: 10.3389/fpls.2021.701406

LEAFY, a Pioneer Transcription Factor in Plants: A Mini-Review

*Nobutoshi Yamaguchi**

Division of Biological Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Nara, Japan

A subset of eukaryotic transcription factors (TFs) possess the ability to reprogram one cell type into another. Genes important for cellular reprograming are typically located in closed chromatin, which is covered by nucleosomes. Pioneer factors are a special class of TFs that can initially engage their target sites in closed chromatin prior to the engagement with, opening of, or modification of the sites by other factors. Although many pioneer factors are known in animals, a few have been characterized in plants. The TF LEAFY (LFY) acts as a pioneer factor specifying floral fate in *Arabidopsis*. In response to endogenous and environmental cues, plants produce appropriate floral inducers (florigens). During the vegetative phase, LFY is repressed by the TERMINAL FLOWER 1 (TFL1)–FD complex, which functions as a floral inhibitor, or anti-florigen. The florigen FLOWERING LOCUS T (FT) competes with TFL1 to prevent the binding of the FD TF to the *LFY* locus. The resulting FT–FD complex functions as a transient stimulus to activate its targets. Once *LFY* has been transcribed in the appropriate spatiotemporal manner, LFY binds to nucleosomes in closed chromatin regions. Subsequently, LFY opens the chromatin by displacing H1 linker histones and recruiting the SWI/SNF chromatin-remodeling complex. Such local changes permit the binding of other TFs, leading to the expression of the floral meristem identity gene *APETALA1*. This mini-review describes the latest advances in our understanding of the pioneer TF LFY, providing insight into the establishment of gene expression competence through the shaping of the plant epigenetic landscape.

Keywords: ***Arabidopsis thaliana*, chromatin, floral meristem identity, histone, LEAFY, pioneer factor**

OPEN ACCESS

Edited by:
Annette Becker, University of Giessen, Germany

Reviewed by:
Cezary Smaczniak, Humboldt University of Berlin, Germany
Chloe Zubieta, Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Nobutoshi Yamaguchi
nobuy@bs.naist.jp

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

Received: *27 April 2021*
Accepted: *01 June 2021*
Published: *05 July 2021*

Citation:
Yamaguchi N (2021) LEAFY, a Pioneer Transcription Factor in Plants: A Mini-Review. Front. Plant Sci. 12:701406. doi: 10.3389/fpls.2021.701406

INTRODUCTION

A subset of eukaryotic transcription factors (TFs) possess reprograming activity to change one cell type into another (Meshi and Iwabuchi, 1995; Drouin, 2014). During cell fate reprograming in eukaryotes, TFs control gene expression programs to enable the formation of distinct cell types from the same genome. Different gene expression programs are blocked by chromatin-mediated mechanisms. TF-binding sites are often masked by nucleosomes, which play important roles in genome packaging and gene expression. The nucleosome consists of a segment of DNA wound around two copies of four types of histone proteins. Nucleosome positions in the genome determine the accessibility of the DNA to regulatory proteins.

A special class of TFs called pioneer factors can access their target DNA sequences inside nucleosomes, typically in chromatin regions where the presence of linker histones represses transcription (Iwafuchi-Doi and Zaret, 2014, 2016; Soufi et al., 2015; Iwafuchi-Doi, 2019).

The primary functions of pioneer factors are cell fate reprograming and the establishment of competence for changes in cellular fate (Zaret and Carroll, 2011). Notable examples include Sox2 and Oct4, two of the four key TFs that together cause the conversion of mammalian somatic cells into induced pluripotent stem cells (Takahashi and Yamanaka, 2006). Molecular genetic, biochemical, and crystal structural analyses have revealed common features of pioneer factors in animals. A pioneer factor in plants was recently identified by two independent groups (Jin et al., 2021; Lai et al., 2021). In this mini-review, the author describes the latest advances in our understanding of this pioneer factor, LEAFY (LFY).

MASTER REGULATORS ARE POTENTIAL CANDIDATES FOR PIONEER TRANSCRIPTION FACTORS

Pioneer TFs are a special group of master regulators. Although not much is known about pioneer TFs in plants, many master regulators have been already identified. Although the definition of the term master regulator or master regulatory gene has expanded since the late 1970s, the original definition was a "gene that occupies the very top of a regulatory hierarchy" which "by its very definition should not be under the regulatory influence of any other gene" (Ohno, 1979). This definition was later modified to describe the hierarchy of cell fate specification in eukaryotes (Nasmyth and Tatchell, 1980; Lewis, 1985; Hamdi et al., 1987; Herskowitz, 1989; Lewis, 1992). Over the next 20 years, the term master regulator was used for genes or proteins with the ability to convert one cell type into another when misexpressed. Classic examples include the myogenic TF MyoD1 in mouse and the hematopoietic TF SCL in zebrafish (Davis et al., 1987; Porcher et al., 1996; Robb et al., 1996; Gering et al., 1998; Tapscott et al., 1998). The basic helix-loop-helix (bHLH) family TF MyoD1 regulates muscle cell differentiation by inducing cell cycle arrest (Olson et al., 2020). Other examples of pioneer factors are the nuclear factor Y (NF-Y) TFs in mouse, and the TFs Oct3/4, Sox2, Klf4, and c-Myc (collectively called the Yamanaka factors) in human and mouse (Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Oldfield et al., 2014). Overall, many animal pioneer factors play key roles in embryogenesis (Lai et al., 2018). Both master regulators and pioneer factors control cell reprograming; therefore, master regulators encoding TFs could be considered candidate pioneer factors. However, the TF families to which most of the animal pioneer factors belong are absent from plants (Lai et al., 2018).

Many TF genes whose activity is sufficient to re-specify cell fate when overexpressed have been identified in *Arabidopsis thaliana*. For example, the master regulator LEAFY COTYLEDON1 (LEC1) is a NF-Y protein that maintains embryonic cell fate during embryogenesis and prevents premature seed germination (West et al., 1994; Lotan et al., 1998; Lee et al., 2003; Tao et al., 2017). Master regulators in plants are involved in cell fate decisions throughout development. For example, the NAC TF VASCULAR-RELATED NAC DOMAIN7 (VND7) promotes xylem vessel cell differentiation (Kubo et al., 2005). Ectopic *VND7* expression was sufficient to confer xylem character. A few bHLH proteins, such as MUTE and FAMA, drive the sequential steps of stomatal differentiation (Ohashi-Ito and Bergmann, 2006; Pillitteri et al., 2007). Overexpression of *FAMA* specified the identity of stomatal and myrosin cells, while *MUTE* misexpression conferred guard cell fate to leaf epidermal cells (Ohashi-Ito and Bergmann, 2006; Pillitteri et al., 2007; Shirakawa et al., 2014). Overexpression of the APETALA 2 (AP2) family TF gene *PLETHORA2* induced ectopic root formation (Aida et al., 2004; Galinha et al., 2007). MADS-box TFs are the core factors involved in floral organ specification (Coen and Meyerowitz, 1991). When overexpressed, they have the ability to transform one type of organ into another (Riechmann and Meyerowitz, 1997). Overexpression of MADS-domain TFs is sufficient to convert leaves into floral organs (Honma and Goto). Among the MADS-box TFs, APETALA1 (AP1) and SEPALLATA3 (SEP3) were proposed to act as pioneer TFs since they can access closed chromatin (Pajoro et al., 2014). The *LFY* gene encodes a plant-specific helix-turn-helix TF (Weigel et al., 1992; Weigel and Nilsson, 1995; Hamès et al., 2008). Although overexpression of *LFY* alone cannot induce ectopic flower formation and does not alter embryogenesis and root formation, overexpression of *LFY* with *WUSCHEL* (encoding a homeodomain TF that promotes stem cell formation) in root tissues conferred floral fate to root cells (Gallois et al., 2004; Wagner et al., 2004). Furthermore, LFY, together with one of its coactivators, the F-box protein UNUSUAL FLORAL ORGANS, can alter leaf development and produce ectopic floral organs (Parcy et al., 1998; Risseeuw et al., 2013).

Among those master regulators, LEC1, AP1, SEP3, and LFY affect chromatin structure (Lai et al., 2018). LEC1 shows sequence similarity to animal pioneer factors. LEC1 might act in the same way as NF-Y in terms of structure and function. MADS-box genes are detected in many eukaryotes, including plants and animals. However, the functional diversification of MADS-domain TFs in plants is much higher than in animals. Indeed, AP1 and SEP3 interact with chromatin remodelers to open chromatin (Smaczniak et al., 2012). On the other hand, MADS-domain TFs in animals act as settler TFs whose genomic binding is principally governed by proximity to open chromatin (Sherwood et al., 2014). Further analysis is required to understand the precise function of the plant MADS domain in the context of chromatin. LFY was the most well-characterized pioneer factor of all the master regulators (Jin et al., 2021; Lai et al., 2021). LFY is only found in plants (Maizel et al., 2005). These results suggested that plant and animal pioneer factors have both the same and different modes of action in terms of structure and function. Further analysis is required to understand the molecular mechanisms of gene expression regulated by plant pioneer factors and its candidates.

THE BASIS AND VALIDATION OF PIONEER FACTORS

Based on our understanding of animals, pioneer factors are characterized by four major properties (**Figure 1**; Iwafuchi-Doi

and Zaret, 2014, 2016; Soufi et al., 2015; Lai et al., 2018; Iwafuchi-Doi, 2019). The first property is direct binding to a target DNA sequence inside a nucleosome (**Figure 1**). This feature is often examined through electrophoretic mobility shift assays (EMSAs; Fernandez et al., 2019; Jin et al., 2021; Lai et al., 2021) and sequential TF and core histone chromatin immunoprecipitation (ChIP; Desvoyes et al., 2018; Jin et al., 2021) to provide evidence that putative pioneer factors have specific chromatin-binding properties suitable for such activities. To exclude the possible contribution to other factors, *in vitro* and *in vivo* experiments are required.

The second property is the initiation of chromatin remodeling (**Figure 1**). To assess this, the chromatin state of a target DNA sequence must be examined before and after the pioneer factors of interest are expressed. Assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq), DNase I hypersensitive site sequencing (DNase-seq), or MNase digestion coupled with high-throughput sequencing (MNase-seq) are often used for this purpose (Zhang et al., 2012, 2015; Bajic et al., 2018). ATAC-seq and DNase-seq are used to measure chromatin openness, while MNase-seq is employed to analyze nucleosome occupancy. Prior to the binding of pioneer factors, target DNA sequences are closed and nucleosomal. However, a causal relationship between pioneer factor binding and chromatin opening must be proven as: Merely identifying a correlation between the target DNA sequences of TFs and chromatin opening sites does not necessarily imply cause and is insufficient to validate pioneer factor status.

The third property is allowing for other factors to bind the chromatin (**Figure 1**). Pioneer factors open up local chromatin regions, thereby directly or indirectly allowing other factors to bind to their targets. Most other TFs cannot initially access a target DNA sequence inside a nucleosome because they lack secondary protein structures important for the recognition of the nucleosome. These non-pioneer TFs are often located alongside the pioneer factor binding sites during or after pioneer factor binding.

The fourth property is the establishment of competence for cell fate changes (**Figure 1**). This feature is usually analyzed through deletion and ectopic expression of a *TF* gene *in vivo* (Sablowski and Meyerowitz, 1998; Wagner et al., 1999, 2004). The effect of the binding of a TF on DNA accessibility at the target sites inside a nucleosome can then be examined. Since the fourth property of pioneer factors and the definition of master regulators largely overlap, master regulators encoding TFs could be considered candidate pioneer factors. Among the master regulators noted above, only LFY meets all four criteria in plants, so far.

REGULATION OF LEAFY REPRESSION AND ACTIVATION

During the vegetative phase, the regulatory region of the *LFY* gene integrates developmental and environmental cues to determine the timing of *LFY* expression (**Figures 2A,B**; Blázquez et al., 1997, 1998; Blázquez and Weigel, 2000). In addition to the 2.3-kilobase-pair upstream intergenic *LFY* promoter region, which contains distal and proximal elements, the genic region of *LFY* also plays key roles in this integration (**Figure 2A**; Blázquez et al., 1997, 1998; Blázquez and Weigel, 2000; Yamaguchi et al., 2009, 2013, 2016; Wu et al., 2015; Zhu et al., 2020). The precocious expression of *LFY* in plants during the vegetative phase led to premature flower formation (Weigel and Nilsson, 1995). As a result, these plants produced few seeds. Thus, *LFY* expression must be repressed until a specific time point (**Figures 2B,C**).

Later work indicated that, to prevent it from specifying floral fate, *LFY* is repressed by the floral inhibitor

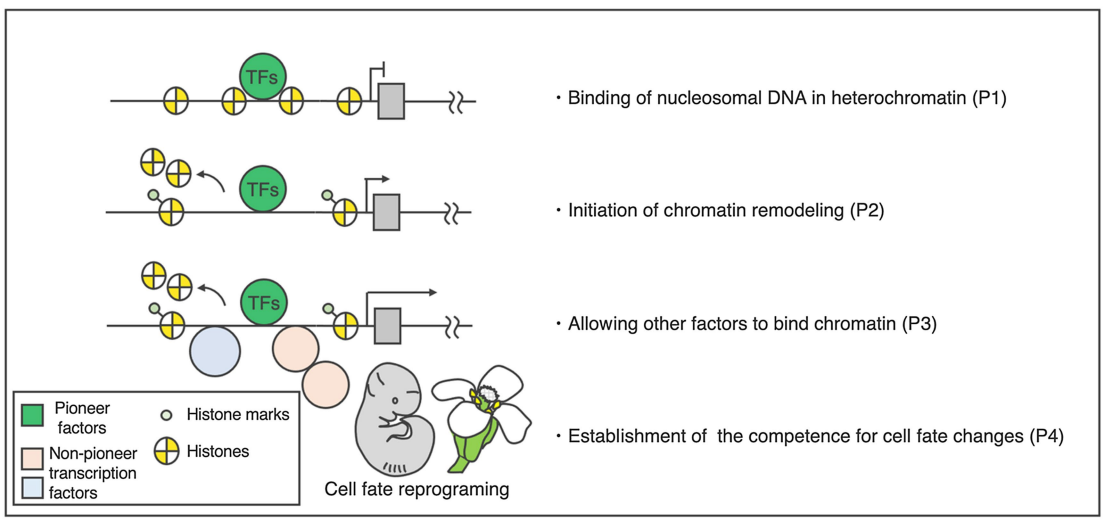

FIGURE 1 | Activity and properties of pioneer factors. Left: hierarchical model of target activation by the pioneer transcription factors. Right: four basic properties of the pioneer factors. The DNA-binding domains of the pioneer factors allow them to target closed chromatin prior to activation [property 1 (P1)]. This binding increases the accessibility of target sites (P2), making the sites accessible to other factors (P3). Pioneer transcription factors play a primary role in cellular programing (P4).

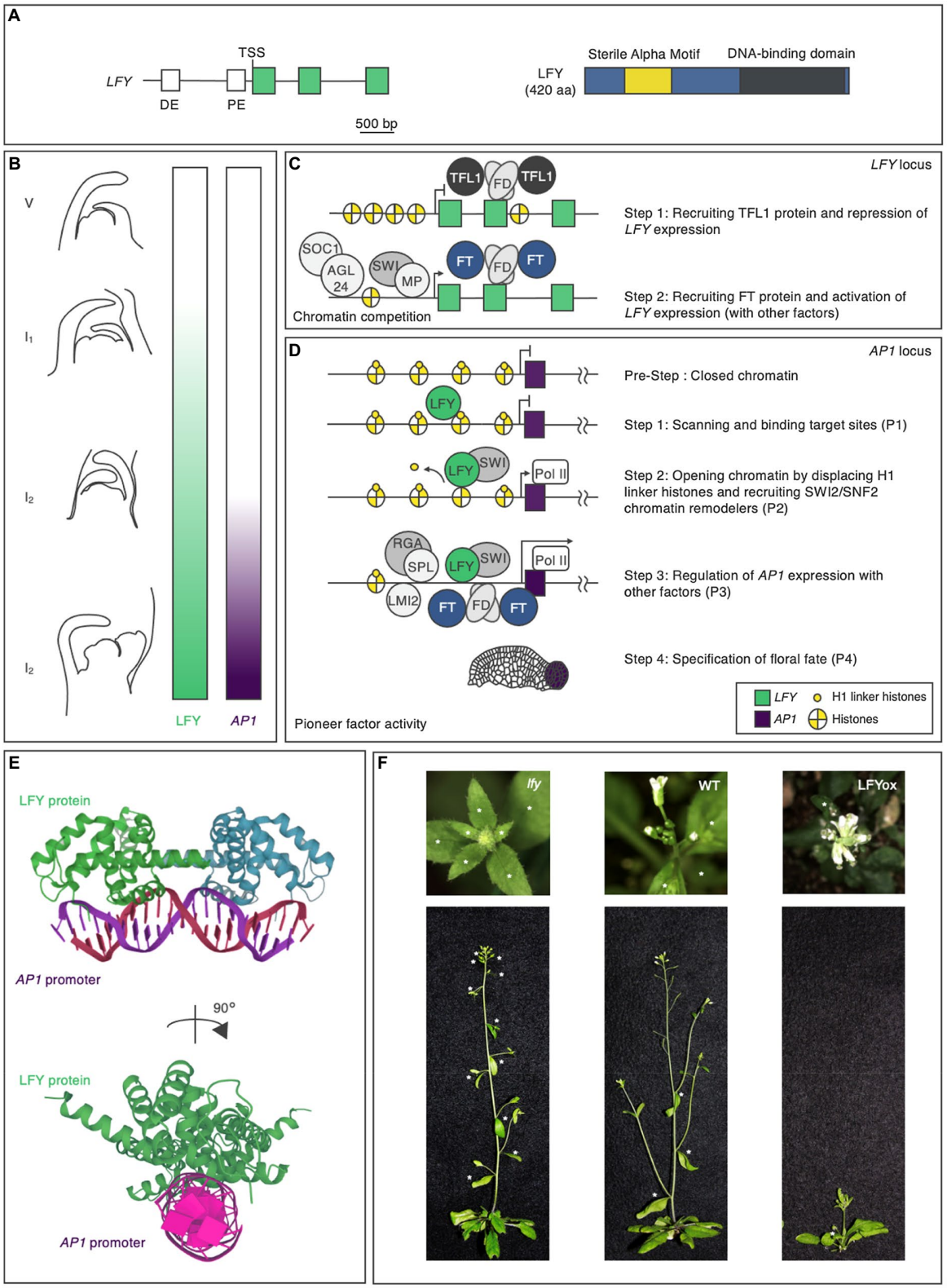

FIGURE 2 | The pioneer transcription factor LEAFY in *Arabidopsis thaliana*. **(A)** *LFY* gene and LFY protein domain structures in *Arabidopsis*. The positions of the conserved regulatory elements (Blázquez and Weigel, 2000) and exons in the gene are indicated in white and green, respectively, and the N-terminal domain,

(Continued)

FIGURE 2 | protein-binding domain, and DNA-binding domain in the protein are indicated in blue, yellow, and dark gray, respectively. DE, distal element; PE, proximal element; and TSS, transcription start site. The scale bar represents nucleotide lengths. **(B)** LFY protein accumulation and *AP1* expression during flower formation. Different lateral organs are formed during each phase of the plant lifecycle. During the vegetative (V) phase, rosette leaves form; during the inflorescence phase 1 (I1), cauline leaves and associated secondary inflorescence branches form; and during the inflorescence phase 2 (I2), flowers form. LFY activity and *AP1* expression are indicated by the green and purple color bars, respectively. **(C)** Hierarchical model of *LFY* activation. **(D)** Hierarchical model of *AP1* activation by the pioneer transcription factor LFY. Properties (P1-P4) are shown in **Figure 1**. **(E)** Crystal structure of LFY and *AP1* double-stranded (ds) DNA. LFY protein is shown in green, while dsDNA is shown in purple. The data were obtained from the Protein Data Bank (https://www.rcsb.org). **(F)** Phenotypes of the *lfy* mutant (left), wild type (middle: WT), and *LFY* overexpressor (right: LFYox). Above: top view. Below: side view. Asterisks indicate secondary inflorescences subtended by cauline leaves on the main stem.

(anti-florigen) TERMINAL FLOWER 1 (TFL1) during the vegetative phase (Bradley et al., 1997; Ratcliffe et al., 1998; Conti and Bradley, 2007). Consistent with the role of TFL1 in repressing *LFY* expression, *tfl1* mutant, and *TFL1* overexpressor plants showed *LFY* overexpressor and *lfy* mutant phenotypes, respectively, in terms of secondary inflorescence number: The knockout mutation of *TFL1* decreased the number of secondary inflorescences, while the constitutive overexpression of *TFL1* increased the number of secondary inflorescences (Ratcliffe et al., 1998). TFL1 is a member of the phosphatidylethanolamine-binding protein (PEBP) family (Jin et al., 2021) and is thought to act as a transcriptional cofactor in a floral repression complex that includes the basic leucine zipper (bZIP) transcription factor FD (Honma and Goto, 2001; Abe et al., 2005; Ho and Weigel, 2014; Collani et al., 2019; Goretti et al., 2020; Zhu et al., 2020). FD recruits TFL1 to the second exon of *LFY* (**Figure 2C**; Zhu et al., 2020). This recruitment is largely dependent on FD activity, as TFL1 occupancy is strongly reduced in the null *fd* mutant background (Zhu et al., 2020). The recruitment of TFL1-FD to *LFY* is mediated by evolutionarily conserved bZIP cis-motifs located at the second exon of this gene (Zhu et al., 2020). As is often the case with exonic TF-binding sites, these sites contribute to *LFY* gene expression.

When conditions are right, plants transition from the vegetative to the reproductive phase. In response to endogenous and environmental cues, plants produce the appropriate floral inducers (florigens). FLOWERING LOCUS T (FT), which (like TFL1) belongs to the PEBP family, acts as a major florigen (Kobayashi et al., 1999; Abe et al., 2005; Wigge et al., 2005). The antagonism between FT and TFL1 involves competition for chromatin-bound FD at the *LFY* locus (Zhu et al., 2020). The resulting FT–FD florigen activation complex functions as a transient stimulus at target loci (**Figure 2C**; Collani et al., 2019; Abe et al., 2019). Other temporal regulators have been identified as *LFY* activators, such as SUPPRESSOR OF CONSTANS OVEREXPRESSION 1, AGAMOUS-LIKE24, SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE3, and MYB33 (**Figure 2C**; Gocal et al., 2001; Liu et al., 2008; Yamaguchi et al., 2009). Whether these activators interact with each other to determine the timing of *LFY* expression in the context of chromatin is unknown.

The timing and the location of *LFY* expression must be specified to confer floral fate on specific cells. The IAA-AUXIN RESPONSE FACTOR (ARF) module and AP2-type TFs control *LFY* expression in floral primordia, as indicated by their similar expression patterns (Karim et al., 2009; Yamaguchi et al., 2013, 2016; Wu et al., 2015). ARF5/MONOPTEROS (MP) and two SWI-SNF ATPase chromatin remodeling factors, SPLAYED (SYD) and BRAHMA (BRM), activate shared targets including *LFY* (**Figure 2C**; Bezhani et al., 2007; Yamaguchi et al., 2013; Wu et al., 2015). The MP-SYD/BRM complex associates with evolutionarily conserved and biologically important auxin response elements (AuxREs) located in the proximal region of the upstream intergenic *LFY* promoter (Yamaguchi et al., 2013; Boer et al., 2014; Wu et al., 2015). The MP-SYD/BRM complex unlocks chromatin and allows shared target loci of AuxREs to become accessible. Numerous genes encoding AP2-type TFs, such as *AINTEGUMENTA* (*ANT*), *AINTEGUMENTA-LIKE 6*/*PLETHORA 3*, *PUCHI*, *DÖRNROSCHEN* (*DRN*), and *DÖRNROSCHEN-LIKE* (*DRNL*), show expression patterns overlapping with that of *LFY* in floral primordia (Nole-Wilson et al., 2005; Karim et al., 2009; Krizek, 2009; Yamaguchi et al., 2013; Chandler and Werr, 2017). In higher-order or sensitized mutants of these genes, *LFY* expression in floral primordia is reduced, pointing to their roles in upregulating *LFY* expression (Yamaguchi et al., 2016). Among these TFs, ANT, and AIL6 moderately bind to the upstream intergenic *LFY* promoter region near the proximal region (Yamaguchi et al., 2016). However, how PUCHI, DRN, and DRNL contribute to the activation of *LFY* expression remains to be clarified.

INITIAL TARGETING OF THE PIONEER FACTOR LEAFY AND SUBSEQUENT EVENTS

Once *LFY* is transcribed in the correct spatiotemporal manner through the actions of the TFs described above, LFY influences fate specification *via* transcriptional regulation, functioning as a pioneer factor. The regulatory network downstream of LFY comprises a set of interlocking feed-forward loops that control the timing of the upregulation of *AP1*, encoding a TF that specifies floral fate (Mandel et al., 1992; Parcy et al., 1998; Wagner et al., 1999; William et al., 2004; Benlloch et al., 2011; Moyroud et al., 2011; Winter et al., 2011; Sayou et al., 2016). LFY specifies not only floral fate, but also flower primordium founder fate and floral organ fate. In this review, the author does not discuss functions other than floral fate specification, since they are covered in detail in recent reviews (Ó'Maoiléidigh, 2014; Wang and Jiao, 2018). LFY meets all four properties for *AP1* regulation (**Figure 2D**).

LFY is composed of two domains, a sterile alpha motif (SAM) oligomerization N-terminal domain and a C-terminal DNA-binding domain (DBD), a helix-turn-helix fold that by itself dimerizes on DNA (**Figure 2A**). Although the SAM oligomerization domain itself does not affect DNA binding *in vitro*, it is required for accessing regions with low-affinity-binding sites and closed chromatin (Sayou et al., 2016). LFY recognizes semi-palindromic 19-bp *cis*-elements through its DBD (**Figure 2E**; Hamès et al., 2008; Moyroud et al., 2011; Winter et al., 2011; Sayou et al., 2016). Besides *cis*-elements in the DNA targets themselves, *in vivo* modifications in the context of chromatin are important for the DNA-binding activity of LFY. EMSA data indicate that LFY associates with the nucleosomal regulatory region of *AP1 in vitro* (Jin et al., 2021; Lai et al., 2021; first property). When the *cis*-element was mutated, LFY did not bind to the nucleosomal substrate, suggesting that LFY binds to nucleosomal DNA *via* its *cis*-element *in vitro* (Jin et al., 2021). MNase-seq and sequential ChIP results supported the notion that LFY binds to the nucleosomal regulatory region of *AP1 in vivo* (Jin et al., 2021).

A LFY-binding test using DAP-seq and ampDAP-seq revealed that LFY is able to bind to both methylated and non-methylated DNA. Whereas an increased number of methylated cytosines in the whole bound region strongly decreases the binding for the two methylation-sensitive TFs (such as ERF018) in DAP relative to ampDAP, LFY binding was only mildly affected (O'Malley et al., 2016; Bartlett et al., 2017; Lai et al., 2021). Since DNA methylation is often seen in closed chromatin regions (Yin et al., 2017; Klemm et al., 2019), pioneer factors may have to access DNA regardless of DNA methylation status. Based on structural analysis of the DBD of LFY in complex with *AP1*, hydrophobic contacts between LFY and DNA could be enhanced by the presence of a methyl group (Hamès et al., 2008). The role of DNA methylation in LFY binding needs to be clarified in the future.

Initial chromatin opening by LFY is mediated by the displacement of the histone H1 linker and the recruitment of SWI/SNF chromatin remodelers (second property). The structural similarity was observed between the helix-turn-helix DBD of LFY and linker histone H1. H1-deficient plants show pleiotropic defects during cell fate specification (Hamès et al., 2008; Rutowicz et al., 2019; Jin et al., 2021). After LFY induction, LFY removes the H1 linker at the *AP1* locus (Jin et al., 2021). LFY interacts with SYD and BRM to open up chromatin by remodeling the nucleosomes at regulatory regions (Bezhani et al., 2007; Wu et al., 2012). SWI3B, a core component of both SYD and BRM, is recruited after LFY induction (Jin et al., 2021). Furthermore, the induction or constitutive expression of *LFY* increases local chromatin accessibility, as revealed by formaldehyde-assisted identification of regulatory elements (FAIRE) analysis (Jin et al., 2021; Lai et al., 2021).

After opening up local chromatin, LFY directly or indirectly allows other TFs to bind their targets (Pastore et al., 2011; Yamaguchi et al., 2014; Jin et al., 2021; third property). In addition to directly activating *AP1*, LFY promotes the activation of regulators of *AP1*. *LATE MERISTEM IDENTITY* (*LMI2*), encoding an MYB TF, is a direct target of LFY. Like LFY, LMI2 also directly promotes *AP1* expression (Pastore et al., 2011). The LFY-binding motif and the LMI2-binding motif in the *AP1* regulatory region are in close proximity (Pastore et al., 2011; Jin et al., 2021). The simultaneous activation of LFY and LMI2 revealed that LMI2 binding in the context of the nucleosome requires the pioneer function of LFY (Jin et al., 2021). LFY also activates the expression of *EUI-LIKE P450 A1* (*ELA1*), which encodes a gibberellin-inactivating enzyme; increased LFY activity leads to reduced gibberellin levels and increased DELLA protein levels (Yamaguchi et al., 2014). A DELLA transcriptional cofactor interacts with the TF SPL9 at the regulatory regions of *AP1* (Yu et al., 2012; Yamaguchi et al., 2014) and activates *AP1* in parallel with LFY. Coherent dual feed-forward loops induce *AP1* expression.

LFY has the ability to convert cell fate when overexpressed (Weigel and Nilsson, 1995; Wagner et al., 1999, 2004; fourth property). Loss or reduction of LFY activity resulted in an increased number of secondary inflorescences, whereas constitutive overexpression of *LFY* caused precocious flower formation without secondary inflorescences (**Figure 2F**; Weigel et al., 1992; Weigel and Nilsson, 1995). LFY conferred floral fate to root explant cells and allowed callus to form flowers and floral organs without producing leaves (Wagner et al., 2004). Regardless of the tissue, LFY alters gene expression programs *via* the same chromatin-mediated mechanisms. Not only in floral cells, but also in root explant cells, the interaction between LFY and nucleosomes, displacement of the histone H1 linker, and recruitment of SWI/SNF chromatin remodelers increased chromatin accessibility, leading to upregulation of *AP1* (Jin et al., 2021; Lai et al., 2021). Since root explants lack floral factors, and the root explants did not previously exhibit floral fate, this indicates that LFY alone is sufficient to trigger cellular reprograming to determine floral fate.

CONCLUSION AND FUTURE PROSPECTS

During cell fate specification in eukaryotes, cellular reprograming is controlled by pioneer TFs. In the past three decades, research into the roles of plant TFs in cell fate specification using phenotypic, transcriptome, epigenome, and crystal structure analyses has revealed the importance of TFs whose misexpression changes the fate of one cell type into another. Two independent groups recently uncovered the initial targeting mechanism by which LFY can engage closed chromatin. This initial targeting of nucleosomal DNA allows LFY to initiate reprograming of silent genes, leading to cell type conversion. Although LFY regulates a lot of downstream target genes involved in flower primordium founder cell fate, and floral organ fate as well, whether LFY also has the potential to function as a pioneer factor in the context of other target genes is not yet known. LFY regulates *AP1* expression as a pioneer factor, but how different target genes are regulated by LFY in a spatiotemporal manner needs further study. There may be differences in the

DNA-binding specificity of LFY in acting as a pioneer factor vs. a non-pioneer factor.

One major limitation to current research on pioneer factors in plants is the lack of a general understanding of these factors; additional pioneer factors in plants need to be identified to provide more data about this class of TFs. Although only LFY fulfills all four criteria of pioneer factors, certainly others will eventually be identified. Although linker histone H1 and SWI/SNF chromatin remodelers play key roles in allowing LFY to exert its roles as a pioneer factor, the initial targeting mechanisms for each pioneer factor could be different. Interestingly, H1- and SWI/SNF-deficient plants show pleiotropic defects, affecting diverse processes including seed dormancy, lateral root formation, root hair fate, stomate formation, and callus formation. There may be pioneer factors that control these developmental processes *via* a mechanism shared with LFY. Various approaches will also be useful for identifying pioneer factors that engage their target sites in chromatin *via* unique mechanisms. Detailed studies of diverse TFs will likely reveal subsets of factors with dominant nucleosome-binding function and pioneer activity in plants. Further understanding of how pioneer factors function will lay the foundation for developing methods to manipulate cell fate in plants.

AUTHOR CONTRIBUTIONS

NY: conceptualization, funding acquisition, and writing.

FUNDING

This work was supported by a grant from a JSPS KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas (no. 18H04782), a JSPS KAKENHI Grant-in-Aid for Scientific Research B (no. 18H02465), a Grant-in-Aid for challenging Exploratory Research (no. 19K22431), and a grant from the SECOM Science and Technology Foundation to NY.

ACKNOWLEDGMENTS

I thank Sachi Ando and Makoto Shirakawa for critical comments and helpful discussion on this manuscript. The author apologized to researchers in the field whose work has not been cited due to space constraints.

REFERENCES

Abe, M., Kobayashi, Y., Yamamoto, S., Daimon, Y., Yamaguchi, A., et al. (2005). FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* 309, 1052–1056. doi: 10.1126/science.1115983

Abe, M., Kosaka, S., Shibuta, M., Nagata, K., Uemura, T., Nakano, A., et al. (2019). Transient activity of the florigen complex during the floral transition in *Arabidopsis thaliana*. *Development* 146:dev171504. doi: 10.1242/dev.171504

Aida, M., Beis, D., Heidstra, R., Willemsen, V., Blilou, I., Galinha, C., et al. (2004). The PLETHORA genes mediate patterning of the *Arabidopsis* root stem cell niche. *Cell* 119, 109–120. doi: 10.1016/j.cell.2004.09.018

Bajic, M., Maher, K. A., and Deal, R. B. (2018). Identification of open chromatin regions in plant genomes using ATAC-seq. *Methods Mol. Biol.* 1675, 183–201. doi: 10.1007/978-1-4939-7318-7_12

Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., et al. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* 12, 1659–1672. doi: 10.1038/nprot.2017.055

Benlloch, R., Kim, M. C., Sayou, C., Thévenon, E., Parcy, F., and Nilsson, O. (2011). Integrating long-day flowering signals: a LEAFY binding site is essential for proper photoperiodic activation of APETALA1. *Plant J.* 67, 1094–1102. doi: 10.1111/j.1365-313X.2011.04660.x

Bezhani, S., Winter, C., Hershman, S., Wagner, J. D., Kennedy, J. F., Kwon, C. S., et al. (2007). Unique, shared, and redundant roles for the *Arabidopsis* SWI/SNF chromatin remodeling ATPases BRAHMA and SPLAYED. *Plant Cell* 19, 403–416. doi: 10.1105/tpc.106.048272

Blázquez, M. A., Green, R., Nilsson, O., Sussman, M. R., and Weigel, D. (1998). Gibberellins promote flowering of *Arabidopsis* by activating the LEAFY promoter. *Plant Cell* 10, 791–800. doi: 10.1105/tpc.10.5.791

Blázquez, M. A., Soowal, L. N., Lee, I., and Weigel, D. (1997). LEAFY expression and flower initiation in *Arabidopsis*. *Development* 124, 3835–3844. doi: 10.1242/dev.124.19.3835

Blázquez, M. A., and Weigel, D. (2000). Integration of floral inductive signals in *Arabidopsis*. *Nature* 404, 889–892. doi: 10.1038/35009125

Boer, D. R., Freire-Rios, A., van den Berg, W. A., Saaki, T., Manfield, I. W., Kepinski, S., et al. (2014). Structural basis for DNA binding specificity by the auxin-dependent ARF transcription factors. *Cell* 156, 577–589. doi: 10.1016/j.cell.2013.12.027

Bradley, D., Ratcliffe, O., Vincent, C., Carpenter, R., and Coen, E. (1997). Inflorescence commitment and architecture in *Arabidopsis*. *Science* 275, 80–83. doi: 10.1126/science.275.5296.80

Chandler, J. W., and Werr, W. (2017). DORNRÖSCHEN, DORNRÖSCHEN-LIKE, and PUCHI redundantly control floral meristem identity and organ initiation in *Arabidopsis*. *J. Exp. Bot.* 68, 3457–3472. doi: 10.1093/jxb/erx208

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

Collani, S., Neumann, M., Yant, L., and Schmid, M. (2019). FT modulates genome-wide DNA-binding of the bZIP transcription factor FD. *Plant Physiol.* 180, 367–380. doi: 10.1104/pp.18.01505

Conti, L., and Bradley, D. (2007). TERMINAL FLOWER1 is a mobile signal controlling *Arabidopsis* architecture. *Plant Cell* 19, 767–778. doi: 10.1105/tpc.106.049767

Davis, R. L., Weintraub, H., and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987–1000. doi: 10.1016/0092-8674(87)90585-x

Desvoyes, B., Sequeira-Mendes, J., Vergara, Z., Madeira, S., and Gutierrez, C. (2018). Sequential ChIP protocol for profiling bivalent epigenetic modifications (ReChIP). *Methods Mol. Biol.* 1675, 83–97. doi: 10.1007/978-1-4939-7318-7_6

Drouin, J. (2014). Minireview: pioneer transcription factors in cell fate specification. *Mol. Endocrinol.* 28, 989–998. doi: 10.1210/me.2014-1084

Fernandez, G. M., Moore, C. D., Schulz, K. N., Alberto, O., Donague, G., Harrison, M. M., et al. (2019). Structural features of transcription factors associating with nucleosome binding. *Mol. Cell* 75, 921–932.E6. doi: 10.1016/j.molcel.2019.06.009

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., et al. (2007). PLETHORA proteins as dose-dependent master regulators of *Arabidopsis* root development. *Nature* 449, 1053–1057. doi: 10.1038/nature06206

Gallois, J. L., Nora, F. R., Mizukami, Y., and Sablowski, R. (2004). WUSCHEL induces shoot stem cell activity and developmental plasticity in the root meristem. *Genes Dev.* 18, 375–380. doi: 10.1101/gad291204

Gering, M., Rodaway, A. R., Göttgens, B., Patient, R. K., and Green, A. R. (1998). The SCL gene specifies haemangioblast development from early mesoderm. *EMBO J.* 17, 4029–4045. doi: 10.1093/emboj/17.14.4029

Gocal, G. F., Sheldon, C. C., Gubler, F., Moritz, T., Bagnall, D. J., Mac Millan, C. P., et al. (2001). GAMYB-like genes, flowering, and gibberellin signaling in *Arabidopsis*. *Plant Physiol.* 127, 1682–1693. doi: 10.1104/pp.010442

Goretti, D., Silvestre, M., Collani, S., Langenecker, T., Méndez, C., Madueño, F., et al. (2020). TERMINAL FLOWER1 functions as a mobile transcriptional cofactor in the shoot apical meristem. *Plant Physiol.* 182, 2081–2095. doi: 10.1104/pp.19.00867

Hamdi, S., Teller, G., and Louis, J. P. (1987). Master regulatory genes, auxin levels, and sexual organogeneses in the dioecious plant *Mercurialis annua*. *Plant Physiol.* 85, 393–399. doi: 10.1104/pp.85.2.393

Hamès, C., Ptchelkine, D., Grimm, C., Thevenon, E., Moyroud, E., Gérard, F., et al. (2008). Structural basis for LEAFY floral switch function and similarity with helix-turn-helix proteins. *EMBO J.* 27, 2628–2637. doi: 10.1038/emboj.2008.184

Herskowitz, I. (1989). A regulatory hierarchy for cell specialization in yeast. *Nature* 342, 749–757. doi: 10.1038/342749a0

Ho, W. W., and Weigel, D. (2014). Structural features determining flower-promoting activity of *Arabidopsis* FLOWERING LOCUS T. *Plant Cell* 26, 552–564. doi: 10.1105/tpc.113.115220

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

Iwafuchi-Doi, M. (2019). The mechanistic basis for chromatin regulation by pioneer transcription factors. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 11:e1427. doi: 10.1002/wsbm.1427

Iwafuchi-Doi, M., and Zaret, K. S. (2014). Pioneer transcription factors in cell reprogramming. *Genes Dev.* 28, 2679–2692. doi: 10.1101/gad.253443.114

Iwafuchi-Doi, M., and Zaret, K. S. (2016). Cell fate control by pioneer transcription factors. *Development* 143, 1833–1837. doi: 10.1242/dev.133900

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

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

Karim, M. R., Hirota, A., Kwiatkowska, D., Tasaka, M., and Aida, M. (2009). A role for *Arabidopsis* PUCHI in floral meristem identity and bract suppression. *Plant Cell* 21, 1360–1372. doi: 10.1105/tpc.109.067025

Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-8

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

Krizek, B. (2009). AINTEGUMENTA and AINTEGUMENTA-LIKE6 act redundantly to regulate *Arabidopsis* floral growth and patterning. *Plant Physiol.* 150, 1916–1929. doi: 10.1104/pp.109.141119

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305

Lai, X., Blanc-Mathieu, R., Grand Vuillemin, L., Huang, Y., Stigliani, A., Lucas, J., et al. (2021). The LEAFY floral regulator displays pioneer transcription factor properties. *Mol. Plant* 14, 829–837. doi: 10.1016/j.molp.2021.03.004

Lai, X., Verhage, L., Hugouvieux, V., and Zubieta, C. (2018). Pioneer factors in animals and plants-colonizing chromatin for gene regulation. *Molecules* 23:1914. doi: 10.3390/molecules23081914

Lee, H., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (2003). *Arabidopsis* LEAFY COTYLEDON1 represents a functionally specialized subunit of the CCAAT binding transcription factor. *Proc. Natl. Acad. Sci. U. S. A.* 100, 2152–2156. doi: 10.1073/pnas.0437909100

Lewis, E. B. (1985). Regulation of the genes of the bithorax complex in drosophila. *Cold Spring Harb. Symp. Quant. Biol.* 50, 155–164. doi: 10.1101/sqb.1985.050.01.021

Lewis, E. B. (1992). The 1991 Albert Lasker medical awards. Clusters of master control genes regulate the development of higher organisms. *JAMA* 267, 1524–1531. doi: 10.1001/jama.1992.03480110100042

Liu, C., Chen, H., Er, H. L., Soo, H. M., Kumar, P. P., Han, J. H., et al. (2008). Direct interaction of AGL24 and SOC1 integrates flowering signals in *Arabidopsis*. *Development* 135, 1481–1491. doi: 10.1242/dev.020255

Lotan, T., Ohto, M., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., et al. (1998). *Arabidopsis* LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93, 1195–1205. doi: 10.1016/S0092-8674(00)81463-4

Maizel, A., Busch, M. A., Tanahashi, T., Perkovic, J., Kato, M., Hasebe, M., et al. (2005). The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science* 308, 260–263. doi: 10.1126/science.1108229

Mandel, M. A., Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1992). Molecular characterization of the *Arabidopsis* floral homeotic gene APETALA1. *Nature* 360, 273–277. doi: 10.1038/360273a0

Meshi, T., and Iwabuchi, M. (1995). Plant transcription factors. *Plant Cell Physiol.* 36, 1405–1420. doi: 10.1093/oxfordjournals.pcp.a078903

Moyroud, E., Minguet, E. G., Ott, F., Yant, L., Posé, D., Monniaux, M., et al. (2011). Prediction of regulatory interactions from genome sequences using a biophysical model for the *Arabidopsis* LEAFY transcription factor. *Plant Cell* 23, 1293–1306. doi: 10.1105/tpc.111.083329

Nasmyth, K. A., and Tatchell, K. (1980). The structure of transposable yeast mating type loci. *Cell* 1980, 753–764. doi: 10.1016/s0092-8674(80)80051-1

Nole-Wilson, S., Tranby, T. L., and Krizek, B. A. (2005). AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol. Biol.* 57, 613–628. doi: 10.1007/s11103-005-0955-6

Ohashi-Ito, K., and Bergmann, D. C. (2006). *Arabidopsis* FAMA controls the final proliferation/differentiation switch during stomatal development. *Plant Cell* 18, 2493–2505. doi: 10.1105/tpc.106.046136

Ohno, S. (1979). Major sex-determining genes. *Monogr. Endocrinol.* 11, 1–140.

Oldfield, A. J., Yang, P., Conway, A. E., Cinghu, S., Freudenberg, J. M., Yellaboina, S., et al. (2014). Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. *Mol. Cell.* 55, 708–722. doi: 10.1016/j.molcel.2014.07.005

Olson, N. J., Fritchie, K. J., Torres-Mora, J., and Folpe, A. L. (2020). MyoD1 expression in fibroepithelial stromal polyps. *Hum. Pathol.* 99, 75–79. doi: 10.1016/j.humpath.2020.03.006

O'Malley, R. C., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., et al. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165, 1280–1292. doi: 10.1016/j.cell.2016.04.038

Ó'Maoiléidigh, D. S., Graciet, E., and Wellmer, F. (2014). Gene networks controlling *Arabidopsis thaliana* flower development. *New Phytol.* 201, 16–30. doi: 10.1111/nph.12444

Pajoro, A., Madrigal, P., Muiño, J. M., Matus, J. T., Jin, J., Mecchia, M. A., et al. (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. *Genome Biol.* 5:R41. doi: 10.1186/gb-2014-15-3-r41

Parcy, F., Nilsson, O., Busch, M. A., Lee, I., and Weigel, D. (1998). A genetic framework for floral patterning. *Nature* 395, 561–566. doi: 10.1038/26903

Pastore, J. J., Limpuangthip, A., Yamaguchi, N., Wu, M. F., Sang, Y., Han, S. K., et al. (2011). LATE MERISTEM IDENTITY2 acts together with LEAFY to activate APETALA1. *Development* 138, 3189–3198. doi: 10.1242/dev.063073

Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L., and Torii, K. U. (2007). Termination of asymmetric cell division and differentiation of stomata. *Nature* 445, 501–505. doi: 10.1038/nature05467

Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F. W., and Orkin, S. H. (1996). The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages. *Cell* 86, 47–57. doi: 10.1016/s0092-8674(00)80076-8

Ratcliffe, O. J., Amaya, I., Vincent, C. A., Rothstein, S., Carpenter, R., Coen, E. S., et al. (1998). A common mechanism controls the life cycle and architecture of plants. *Development* 125, 1609–1615.

Riechmann, J. L., and Meyerowitz, E. M. (1997). Determination of floral organ identity by *Arabidopsis* MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity. *Mol. Biol. Cell* 8, 1243–1259. doi: 10.1091/mbc.8.7.1243

Risseeuw, E., Venglat, P., Xiang, D., Komendant, K., Daskalchuk, T., Babic, V., et al. (2013). An activated form of UFO alters leaf development and produces ectopic floral and inflorescence meristems. *PLoS One* 8:e83807. doi: 10.1371/journal.pone.0083807

Robb, L., Elwood, N. J., Elefanty, A. G., Köntgen, F., Li, R., Barnett, L. D., et al. (1996). The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse. *EMBO J.* 15, 4123–4129.

Rutowicz, K., Lirski, M., Mermaz, B., Teano, G., Schubert, J., Mestiri, I., et al. (2019). Linker histones are fine-scale chromatin architects modulating developmental decisions in *Arabidopsis*. *Genome Biol.* 20:157. doi: 10.1186/s13059-019-1767-3

Sablowski, R. W., and Meyerowitz, E. M. (1998). A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. *Cell* 92, 93–103. doi: 10.1016/s0092-8674(00)80902-2

Sayou, C., Nanao, M. H., Jamin, M., Posé, D., Thévenon, E., Grégoire, L., et al. (2016). A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat. Commun.* 7:11222. doi: 10.1038/ncomms11222

Sherwood, R. I., Hashimoto, T., O'Donnell, C. W., Lewis, S., Barkal, A. A., van Hoff, J. P., et al. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* 32, 171–178. doi: 10.1038/nbt.2798

Shirakawa, M., Ueda, H., Nagano, A. J., Shimada, T., Kohchi, T., and Hara-Nishimura, I. (2014). FAMA is an essential component for the differentiation of two distinct cell types, myrosin cells and guard cells, in Arabidopsis. *Plant Cell* 26, 4039–4052. doi: 10.1105/tpc.114.129874

Smaczniak, C., Immink, R. G., Muiño, J. M., Blanvillain, R., Busscher, M., Busscher-Lange, J., et al. (2012). Characterization of MADS-domain transcription factor complexes in *Arabidopsis* flower development. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1560–1565. doi: 10.1073/pnas.1112871109

Soufi, A., Garcia, M. F., Jaroszewicz, A., Osman, N., Pellegrini, M., and Zaret, K. S. (2015). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 161, 555–568. doi: 10.1016/j.cell.2015.03.017

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., et al. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872. doi: 10.1016/j.cell.2007.11.019

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676. doi: 10.1016/j.cell.2006.07.024

Tao, Z., Shen, L., Gu, X., Wang, Y., Yu, H., and He, Y. (2017). Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature* 551, 124–128. doi: 10.1038/nature24300

Tapscott, S. J., Davis, R. L., Thayer, M. J., Cheng, P. F., Weintraub, H., and Lassar, A. B. (1998). MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* 242, 405–411. doi: 10.1126/science.3175662

Wagner, D., Sablowski, R. W., and Meyerowitz, E. M. (1999). Transcriptional activation of APETALA1 by LEAFY. *Science* 285, 582–584. doi: 10.1126/science.285.5427.582

Wagner, D., Wellmer, F., Dilks, K., William, D., Smith, M. R., Kumar, P. P., et al. (2004). Floral induction in tissue culture: a system for the analysis of LEAFY-dependent gene regulation. *Plant J.* 39, 273–282. doi: 10.1111/j.1365-313X.2004.02127.x

Wang, Y., and Jiao, Y. (2018). Auxin and above-ground meristems. *J. Exp. Bot.* 69, 147–154. doi: 10.1093/jxb/erx299

Weigel, D., Alvarez, J., Smyth, D. R., Yanofsky, M. F., and Meyerowitz, E. M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* 69, 843–859. doi: 10.1016/0092-8674(92)90295-n

Weigel, D., and Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* 377, 495–500. doi: 10.1038/377495a0

West, M., Yee, K. M., Danao, J., Zimmerman, J. L., Fischer, R. L., Goldberg, R. B., et al. (1994). LEAFY COTYLEDON1 is an essential regulator of late embryogenesis and cotyledon identity in *Arabidopsis*. *Plant Cell* 6, 1731–1745. doi: 10.2307/3869904

Wigge, P. A., Kim, M. C., Jaeger, K. E., Busch, W., Schmid, M., Lohmann, J. U., et al. (2005). Integration of spatial and temporal information during floral induction in Arabidopsis. *Science* 309, 1056–1059. doi: 10.1126/science.1114358

William, D. A., Su, Y., Smith, M. R., Lu, M., Baldwin, D. A., and Wagner, D. (2004). Genomic identification of direct target genes of LEAFY. *Proc. Natl. Acad. Sci. U. S. A.* 101, 1775–1780. doi: 10.1073/pnas.0307842100

Winter, C. M., Austin, R. S., Blanvillain-Baufumé, S., Reback, M. A., Monniaux, M., Wu, M. F., et al. (2011). LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Dev. Cell* 20, 430–443. doi: 10.1016/j.devcel.2011.03.019

Wu, M. F., Sang, Y., Bezhani, S., Yamaguchi, N., Han, S. K., Li, Z., et al. (2012). SWI2/SNF2 chromatin remodeling ATPases overcome polycomb repression and control floral organ identity with the LEAFY and SEPALLATA3 transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3576–3581. doi: 10.1073/pnas.1113409109

Wu, M. F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., et al. (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *Life* 4:e09269. doi: 10.7554/eLife.09269

Yamaguchi, N., Jeong, C. W., Nole-Wilson, S., Krizek, B. A., and Wagner, D. (2016). AINTEGUMENTA and AINTEGUMENTA-LIKE6/PLETHORA3 induce LEAFY expression in response to auxin to promote the onset of flower formation in *Arabidopsis*. *Plant Physiol.* 170, 283–293. doi: 10.1104/pp.15.00969

Yamaguchi, N., Winter, C. M., Wu, M. F., Kanno, Y., Yamaguchi, A., Seo, M., et al. (2014). Gibberellin acts positively then negatively to control onset of flower formation in *Arabidopsis*. *Science* 344, 638–641. doi: 10.1126/science.1250498

Yamaguchi, N., Wu, M. F., Winter, C. M., Berns, M. C., Nole-Wilson, S., Yamaguchi, A., et al. (2013). A molecular framework for auxin-mediated initiation of flower primordia. *Dev. Cell* 24, 271–282. doi: 10.1016/j.devcel.2012.12.017

Yamaguchi, A., Wu, M. F., Yang, L., Wu, G., Poethig, R. S., and Wagner, D. (2009). The micro RNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1. *Dev. Cell* 17, 268–278. doi: 10.1016/j.devcel.2009.06.007

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* 356:eaaj2239. doi: 10.1126/science.aaj2239

Yu, S., Galvão, V. C., Zhang, Y. C., Horrer, D., Zhang, T. Q., Hao, Y. H., et al. (2012). Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. *Plant Cell* 24, 3320–3332. doi: 10.1105/tpc.112.101014

Zaret, K. S., and Carroll, J. S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241. doi: 10.1101/gad.176826.111

Zhang, T., Zhang, W., and Jiang, J. (2015). Genome-wide nucleosome occupancy and positioning and their impact on gene expression and evolution in plants. *Plant Physiol.* 168, 1406–1416. doi: 10.1104/pp.15.00125

Zhang, W., Zhang, T., Wu, Y., and Jiang, J. (2012). Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. *Plant Cell* 24, 2719–2731. doi: 10.1105/tpc.112.098061

Zhu, Y., Klasfeld, S., Jeong, C. W., Jin, R., Goto, K., Yamaguchi, N., et al. (2020). TERMINAL FLOWER 1-FD complex target genes and competition with FLOWERING LOCUS T. *Nat. Commun.* 11:5118. doi: 10.1038/s41467-020-18782-1

Conflict of Interest: The author declares 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 in Plant Science

ORIGINAL RESEARCH
published: 21 July 2021
doi: 10.3389/fpls.2021.708573

MIR156-Targeted SPL9 Is Phosphorylated by SnRK2s and Interacts With ABI5 to Enhance ABA Responses in *Arabidopsis*

Huixue Dong[1†], Suli Yan[1†], Yexing Jing[1], Ruizhen Yang[1], Yunwei Zhang[1], Yun Zhou[2], Yingfang Zhu[2] and Jiaqiang Sun[1]*

[1] National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [2] State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng, China

The miR156-targeted SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factors play key roles in regulating plant development, but little is known about their function in abscisic acid (ABA) signaling. Here, we report that the miR156-targeted SPLs enhance ABA responses and contribute to the inhibition of pre-harvest sprouting. We find that SPL9 directly activates the expression of ABA responsive genes through binding to their promoters. SPL9 was further shown to physically interact with ABSCISIC ACID INSENSITIVE 5 (ABI5), a master transcription factor in ABA signaling, thus promoting its association with the promoters of ABA responsive genes. Furthermore, we reveal that the protein kinases SnRK2s interact with and phosphorylate SPL9, which is essential for its role in the activation of ABA responses. Together, our results disclose a SnRK2s-SPLs-ABI5 regulatory module in ABA signaling in *Arabidopsis*.

Keywords: miR156, SPLs, ABA, ABI5, *Arabidopsis*

OPEN ACCESS

Edited by:
Gang Wu,
Zhejiang Agriculture and Forestry University, China

Reviewed by:
Kewei Zhang,
Zhejiang Normal University, China
Kun-Ming Chen,
Northwest A&F University, China

***Correspondence:**
Jiaqiang Sun
sunjiaqiang@caas.cn

[†] These authors have contributed equally to this work

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

Received: *12 May 2021*
Accepted: *26 May 2021*
Published: *21 July 2021*

Citation:
Dong H, Yan S, Jing Y, Yang R, Zhang Y, Zhou Y, Zhu Y and Sun J (2021) MIR156-Targeted SPL9 Is Phosphorylated by SnRK2s and Interacts With ABI5 to Enhance ABA Responses in Arabidopsis. Front. Plant Sci. 12:708573. doi: 10.3389/fpls.2021.708573

INTRODUCTION

The stress-related phytohormone abscisic acid (ABA) inhibits seed germination and seedling growth to adapt various environmental challenges (Cutler et al., 2010; Weiner et al., 2010). Molecular genetics studies have significantly advanced our understanding on the molecular basis of ABA signaling in *Arabidopsis*. Among them, ABA-INSENSITIVE1 (ABI1) (Leung et al., 1994; Gosti et al., 1999) and ABI2 (Leung et al., 1997; Rodriguez et al., 1998) are clade A protein phosphatase 2Cs (PP2Cs), which negatively regulate ABA signaling during seed germination. However, the downstream B3 transcription factor ABI3, AP2 transcription factor ABI4 and bZIP transcription factor ABI5 positively regulate the ABA-inhibited seed germination and early seedling development (Giraudat et al., 1992; Finkelstein et al., 1998, 2011; Finkelstein and Lynch, 2000; Umezawa et al., 2010). Several SNF1 (sucrose non-fermenting 1)-related kinase 2s (SnRK2s), including SnRK2.2, SnRK2.3, and SnRK2.6 (also known as Open Stomata 1, OST1), were identified as stress- or ABA-activated protein kinases and function redundantly in ABA-mediated regulation of seed germination, seedling growth, drought stress and stomatal closure (Mustilli et al., 2002; Fujii et al., 2007; Nakashima et al., 2009; Ding et al., 2015; Huang et al., 2018).

Since the identification of ABA receptors, PYRABACTIN RESISTANCE1 (PYR1)/PYR1-LIKE (PYL)/REGULATORY COMPONENTS OF ABA RECEPTORS (RCAR) (Fujii et al., 2009; Melcher et al., 2009; Miyazono et al., 2009; Santiago et al., 2009; Soon et al., 2012), a core ABA signaling pathway has been discovered. In the absence of ABA, PP2Cs inhibit the activity of SnRK2s by physical interaction and dephosphorylation (Fujii et al., 2009; Ma et al., 2009; Umezawa et al., 2009), leading to inhibition of downstream transcription factors required for ABA-responsive gene expression (Kobayashi et al., 2005). Perception of ABA by its receptors PYR/PYL/RCAR, facilitates the interaction between PYR/PYL/RCAR and PP2Cs to prevent PP2Cs inhibition on SnRK2s activity (Fujii et al., 2009; Ma et al., 2009; Park et al., 2009). Thus, the ABA-activated SnRK2s phosphorylate and activate the downstream transcription factors (e.g., ABI5) to regulate ABA responsive gene expression (Kobayashi et al., 2005; Fujii et al., 2007; Nakashima et al., 2009).

The SQUAMOSA PROMOTER BINDING PROTEIN (SBP)-like (SPL) belongs to plant-specific transcription factors and contains a highly conserved SBP-box domain (Cardon et al., 1999), which was revealed to specifically bind the core cis-element GTAC (Yamasaki et al., 2004; Birkenbihl et al., 2005; Liang et al., 2008; Lu et al., 2013). In *Arabidopsis*, *SPL* genes are divided into two subgroups, represented by *SPL3* (including *SPL3*, *SPL4*, and *SPL5*) which encodes a small protein, and *SPL9* (including *SPL2*, *SPL6*, *SPL9*, *SPL10*, *SPL11*, *SPL13*, and *SPL15*) which encodes a much larger protein, respectively (Cardon et al., 1999; Yang et al., 2008). Among them, some *SPL* genes such as *SPL3*, *SPL9*, and *SPL15* are regulated by microRNA156 (miR156) (Schwab et al., 2005; Wu and Poethig, 2006; Xing et al., 2010). The miR156-targeted SPL transcription factors play key roles in plant growth and development. For example, *SPL3*, *SPL4*, *SPL5,* and *SPL9* function in the control of flowering time and phase transition (Wu and Poethig, 2006; Wang et al., 2009); *SPL9* and its paralog *SPL15* regulate shoot branching (Schwarz et al., 2008). In addition, recent studies reported that overexpression or knockdown of miR156 can affect seed germination and dormancy in *Arabidopsis* and rice (Huo et al., 2016; Liu et al., 2019; Miao et al., 2019).

In this study, we uncover that miR156-targeted SPLs transcription factors positively regulate ABA responses and inhibit pre-harvest sprouting (PHS) in *Arabidopsis*. We demonstrate that SPLs interact with the master transcription factor ABI5 to promote ABA signaling. Furthermore, we show that SnRK2s physically interact with and phosphorylate SPLs. Importantly, the ABA-induced SPL9 phosphorylation is required for its function in the activation of ABA responses.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

Arabidopsis thaliana ecotype Col-0 was used as the wild type. Some of the plant materials used in this study were previously described: *GFP-rSPL9* (Wang et al., 2009); *rSPL3-HA* (Wang et al., 2009); *MIM156* (Wang et al., 2009); *MIR156* (Xie et al., 2017); *abi5-7* (Chen et al., 2012), and *snrk2.2/2.3/2.6* (Fujii and Zhu, 2009). The *GFP-rSPL9/abi5-7* and *GFP-rSPL9/snrk2.2/2.3/2.6* lines were generated by genetic crossing between *GFP-rSPL9* and *abi5-7* or *snrk2.2/2.3/2.6*, respectively.

Arabidopsis seedlings were grown on half-strength Murashige and Skoog (MS) solid medium containing 2% sucrose at 22°C in a light incubator with 16-h-light/8-h-dark photoperiod. *N. benthamiana* plants were grown under a 16-h-light/8-h-dark cycle in a greenhouse at 22°C for 1 month before infiltration.

DNA Constructs and Transgenic Plants

For BiFC assays, gateway cloning strategy (Invitrogen) was used. The full-length coding sequence (CDS) of *SPL9* or *SPL3* was cloned into *pQBV3* vector (Dong et al., 2020) and subsequently introduced into the destination vector *pEarleygate202-YN* (cYFP) (Lu et al., 2010). Similarly, the full-length CDS of *ABI5* was introduced into the *pEarleygate201-YN* (nYFP) vector (Lu et al., 2010).

For LCI assays, the full-length CDS of *SPL9* was cloned into *p1300-35S-nLUC* vector or *p1300-35S-cLUC* vector (Chen et al., 2008) to generate nLUC-SPL9 or cLUC-SPL9. Similarly, the CDSs of *ABI3*, *ABI4*, *ABI5*, and *SnRK2s* were cloned into *p1300-35S-nLUC* vector or *p1300-35S-cLUC* vector (Chen et al., 2008), respectively. The truncated versions of *SPL9* or *ABI5* were cloned into *p1300-35S-cLUC* vector (Chen et al., 2008), respectively.

For pull-down assays, the full-length CDS of *SPL9* or *SPL3* was inserted into *pMAL-c2X* vector to generate MBP-SPL9 and MBP-SPL3, respectively. Then, the MBP-SPL9 construct was mutated to MBP-SPL9(2A) using the Site-Directed Mutagenesis Kit (Mei5 Biotechnology, MF129-01). Similarly, the full-length CDS of *ABI5* was inserted into *pGEX4T-1* vector to generate ABI5-GST. All the ligations above were performed based on ligation free cloning master mix (Applied Biological Materials, E011-5-A) according to the manufacturer's instruction.

For *Em6*$_{pro}$*:LUC* and *Em1*$_{pro}$*:LUC* constructs, the ∼1.5-kb promoter of *Em6* and 800-bp promoter of *Em1* were separately ligated into the entry vector *pQBV3*, and then introduced into the vector *pGWB35* (Nakagawa et al., 2007). The construct of *35S:rSPL9-MYC* was generated based on the destination vector *pGWB17* (Nakagawa et al., 2007). The constructs of *rSPL9-YFP* and *rSPL9(2A)-YFP* were generated based on the destination vector *pEarly-101* driven by the 35S promoter.

To generate the *SPL9*$_{pro}$*:GFP-rSPL9*, *SPL9*$_{pro}$*:GFP-rSPL9(2A)* and *GFP-rSPL9/SnRK2.6-Flag* transgenic plants, 2-kb promoter of *SPL9* was ligated into *p1305-35S-GFP* to produce *p1305-SPL9*$_{pro}$*-GFP*, next the full length CDS of *SPL9* or *SPL9(2A)* was introduced in it to generate *SPL9*$_{pro}$*:GFP-rSPL9* or *SPL9*$_{pro}$*:GFP-rSPL9(2A)* construct, respectively. *SnRK2.6* gene was amplified and inserted into the *p1300-35S-Flag* vector. *Agrobacterium* strain GV3101 carrying the construct was then transformed into the Col-0 or *GFP-rSPL9* plants to generate *SPL9*$_{pro}$*:GFP-rSPL9*, *SPL9*$_{pro}$*:GFP-rSPL9(2A)* or *GFP-rSPL9/SnRK2.6-Flag* transgenic plants using the floral-dip method (Clough and Bent, 1998), respectively.

All the primers used for the constructs above are summarized in **Supplementary Table 1** and the constructs described above are summarized in **Supplementary Table 2**.

RNA Extraction and Gene Expression Analyses

Total RNA was extracted using Trizol (Invitrogen) reagent according to the manufacture's instruction. About 2 μg of total RNA were used for reverse transcription with the 5× All-In One RT MasterMix system (Applied Biological Materials). Quantitative real-time polymerase chain reaction (qRT-PCR) assay was performed using SYBR® Premix Ex Taq Kit (TaKaRa), and the expression levels of *ACT7* were used as the internal control. The primer sequences are listed in **Supplementary Table 3**.

ABA Treatment Assays and Pre-Harvest Sprouting

For ABA responses, seeds of different genotypes were harvested at the same time for the germination and cotyledon greening assays as described before (Bu et al., 2009; Li et al., 2011). Seeds of different genotypes were sown on the same 1/2 MS medium supplemented with different ABA concentrations as indicated and chilled at 4°C in the dark for 2 days (stratified). Then the seeds were moved to 22°C with a 16-h-light/8-h-dark cycle in a light chamber. The percentage of seed germination or cotyledon greening was scored at 3 or 5 days after the end of stratification, respectively. Germination was defined as an obvious emergence of the radicle through the seed coat. Cotyledon greening is defined as obvious cotyledon expansion and turning green (Bu et al., 2009; Li et al., 2011; Chen et al., 2012). For the PHS test, plants with early siliques that matured at the same time were directly sown on water saturated filter paper then placed in the growth chamber with a 16-h-light/8-h-dark cycle.

Firefly Luciferase Complementation Imaging (LCI) Assays

The luciferase complementation imaging (LCI) assays for the protein interaction detection was performed in *N. benthamiana* leaves as described previously (Chen et al., 2008). The indicated genes were fused into nLUC or cLUC, respectively, and separately introduced into *Agrobacterium* strain GV3101. Then, *Agrobacteria* cells carrying nLUC or cLUC derivative constructs were co-injected in *N. benthamiana* leaves. The LUC activities were analyzed using NightSHADE LB 985 (Berthold).

Chromatin Immunoprecipitation-qPCR Assays

The 6-day-old *Arabidopsis* seedlings grown on 1/2 MS medium were treated with or without 50 μM ABA for 2 h and then collected for chromatin immunoprecipitation (ChIP) assays as previously described (Zhu et al., 2012). Briefly, about 2 to 3 grams of each sample were cross-linked in 1% formaldehyde under vacuum for 15 min, followed by 5-min neutralization with 0.125 M glycine. The samples were separately immunoprecipitated with or without anti-GFP antibody (Abcam, ab290). Finally, the GFP-specific enrichment of the fragments from *Em1* or *Em6* promoter was analyzed by qPCR using specific primer sets listed in **Supplementary Table 4**. The enrichment fold of a certain fragment was calculated by normalizing to the amount of no antibody-immunoprecipitates DNA samples.

Subcellular Localization and Bimolecular Fluorescence Complementation (BiFC) Assays

For localization experiments, *Agrobacterium* GV3101 harboring the *rSPL9-YFP* or *rSPL9(2A)-YFP* construct was injected into *N. benthamiana* leaves. For BiFC assays, the indicated vectors were co-transformed into *Agrobacterium* GV3101 and then co-expressed in *N. benthamiana* leaves as described previously (Dong et al., 2020). The injected tobacco leaves were incubated for 48 h, and then the fluorescence signal of yellow fluorescent protein (YFP) was observed using the confocal microscope (Carl Zeiss, LSM880).

Protein Extraction, Immunoblotting, and Co-immunoprecipitation (Co-IP) Analyses

The GFP-SPL9 fusion proteins were extracted from the 6-day-old *GFP-rSPL9* or *GFP-rSPL9(2A)* transgenic plants using the extracted buffer (125 mM Tris–HCl [pH 6.8], 4% SDS, 20% glycerol, 0.001% Bromophenol blue, 2% β–Mercaptoethanol). For the immunoblotting detection of GFP-SPL9, we used anti-GFP (1:2000; Roche, 11814460001) antibody. ACT (1:5000; CWBIO, CW0264) was employed as a loading control.

The Col-0, *GFP-rSPL9* transgenic plants and anti-ABI5 antibody were used in the Co-IP assays for the interaction of SPL9 and ABI5. Total proteins were extracted from the 6-day-old seedlings treated with 50 μM ABA for 2 h using the lysis buffer (50 mM Tris–HCl [pH 7.5], 150 mM NaCl, 5 mM EDTA [pH 8.0], 0.2% Triton X-100, 0.2% NP-40, 20 μM MG132) with freshly added PMSF (0.6 mM) and 1× protease inhibitor. The extracts were centrifuged for 20 min and the supernatant was incubated with anti-GFP magnetic beads (MBL, D153-10) overnight. Next, the beads were washed five times with the lysis buffer and eluted samples were analyzed by immunoblotting with anti-GFP and anti-ABI5 (1:5000; Agrisera, AS121863) antibodies.

The 6-day-old *GFP-rSPL9* and *GFP-rSPL9/SnRK2.6-Flag* transgenic plants treated with 50 μM ABA plus 30 μM MG132 for 4 h were used in the Co-IP assays for the interaction of SnRK2.6 and SPL9. Total proteins were extracted as described above. The supernatant was incubated with anti-Flag magnetic beads (MBL, M185-10) overnight. Proteins were detected with anti-GFP and anti-Flag (1:5000; MBL, M185-3L) antibodies, respectively.

In vitro and Semi-*in vivo* Pull-Down Assays

The constructs (MBP, MBP-SPL9, MBP-SPL3, GST, and ABI5-GST) were separately transformed into *Escherichia coli* transetta. The fusion proteins were induced with 4 mM isopropyl β-D-thiogalactopyranoside (IPTG) at 18°C overnight. For the pull-down assays of SPL9 and ABI5, the fusion proteins were incubated with glutathione resin (GenScript) overnight in column buffer (20 mM Tris–HCl [pH 7.5], 200 mM NaCl,

1 mM PMSF, 1 mM DTT and 1× protease inhibitor (Roche 4693132001)]. For the pull-down assays of SPL3 and ABI5, the fusion proteins were incubated with amylose resin (New England Biolabs) overnight in column buffer. Next, the GST bind resin or MBP bind resin was washed five times with column buffer, resolved by SDS-PAGE, and detected using anti-GST (1:3000, CW0144, CWbiotech) and anti-MBP (1:3000, CW0288, CWbiotech) antibodies.

The 6-day-old *GFP-rSPL9* seedlings treated with 50 μM ABA for 2 h and SnRK2.6-His fusion proteins were used for the semi-*in vivo* pull-down assays. The GFP-SPL9 proteins were extracted with lysis buffer with freshly added PMSF (0.6 mM) and 1× protease inhibitor. Then SnRK2.6-His fusion proteins were incubated with the GFP-SPL9 protein extracts overnight and added Ni-NTA resin (TransGen Biotech, DP101-01) for a further 2 h. The His bind resin was washed five times with PBS buffer (CWBIO, CW0040S), resolved by SDS-PAGE, and detected using anti-His (1:3000; CWBIO, CW0143M) and anti-GFP (1:2000; Roche, 11814460001) antibodies, respectively.

In vitro and *in vivo* Phosphorylation Assays

For the *in vitro* phosphorylation assays, 1 μg MBP-SPL9, MBP-SPL9(2A) or MBP-SPL3 fusion proteins were incubated with 1 μg SnRK2.6-His in 20 μl kinase reaction buffer (25 mM Tris–HCl [pH 7.5], 12 mM $MgCl_2$, 1 mM DTT and 1 mM ATP) at 37°C for 1 h. The reactions were boiled with 5× SDS loading buffer then separated by phos-tag SDS-PAGE (Kinoshita et al., 2006). The signals were detected using anti-MBP antibody.

For the *in vivo* kinase assays, the GFP-SPL9 fusion proteins were extracted with buffer (150 mM KCl, 50 mM HEPES [PH7.5], 0.4% Triton X-100, 1 mM DTT, 1× protease inhibitor and phosphatase inhibitor cocktail) and immunoprecipitated with anti-GFP magnetic beads. Then the IP products were separated by phos-tag SDS-PAGE and analyzed with anti-GFP antibody.

Transcriptional Activity Assays in *N. benthamiana*

The transcriptional activity assays were carried out in *N. benthamiana* leaves as previously described (Sun et al., 2012). In brief, the reporter Em_{pro}*:LUC* and effector *35S:rSPL9-MYC* were separately introduced into *Agrobacterium* GV3101 to perform the con-infiltration in *N. benthamiana* leaves. The *N. benthamiana* leaves after infiltrating 24 h were injected with 50 μM ABA and incubated for a further 24 h. The luciferase luminescence was observed using NightSHADE LB 985 (Berthold), and quantification of luciferase activities were carried out with IndiGO software (version 2.03.0).

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: *SPL9* (At2g42200), *SPL3* (At2g33810), *ABI5* (At2g36270), *ABI3* (At3g24650), *ABI4* (At2g40220), *SnRK2.2* (AT3G50500), *SnRK2.3* (AT5G66880), *SnRK2.6* (AT4G33950), *Em1* (AT3G51810), and *Em6* (AT2G40170).

RESULTS

The miR156-Targeted SPLs Enhance ABA Responses

To investigate a potential role of the miR156-regulated SPLs in the ABA signaling, we tested the seed germination and cotyledon greening phenotypes of SPLs-related transgenic lines in response to ABA. The *GFP-rSPL9* line is identical to a gain-of-function mutant of *SPL9* gene, in which a miR156-resistant version of *SPL9* is expressed from its native promoter (Wang et al., 2009), and the *MIM156* line has elevated expression of *SPL9* and other *SPLs* (Wang et al., 2009). In the absence of exogenously supplied ABA, the seed germination and cotyledon greening percentages of different genotypes were comparable (**Figure 1A** and **Supplementary Figure 1**). However, the seed germination and cotyledon greening of *GFP-rSPL9* and *MIM156* seedlings were much lower than the wild-type Columbia-0 (Col-0) under ABA treatment (**Figure 1A** and **Supplementary Figure 1**), indicating that overexpression of SPL9 and SPL3 conferred ABA hypersensitivity. Thus, the miR156-regulated SPL9 appears to play a positive role in regulating ABA responses. Meanwhile, SPL3 also positively regulates ABA responses in seed germination and cotyledon greening (**Supplementary Figure 1**). These results suggest that the miR156-targeted SPLs play an enhancing effect on the ABA response during seed germination and early seedling development.

Considering that ABA plays a critical role in preventing PHS, which occurs when adequate temperature and humid conditions prevail during late maturation of crops in the field, we wondered whether the miR156-SPL9 module plays a role in preventing PHS. We conducted germination assays using mature siliques of the Col-0, *GFP-rSPL9* and *MIM156* plants. Interestingly, the freshly harvested seeds from unopened siliques of *GFP-rSPL9* and *MIM156* displayed greatly increased dormancy compared with Col-0 (**Figure 1B**).

We further investigated whether the miR156-targeted SPLs regulate the transcriptional expression of ABA responsive genes. Quantitative reverse transcriptase-PCR (qRT-PCR) analyses showed that ABA-induced expression levels of the representative ABA-responsive genes *Em1* and *Em6* were dramatically enhanced in the *GFP-rSPL9* and *MIM156* seedlings compared with Col-0 plants (**Figure 1C**). In contrast, the ABA-induced expression levels of *Em6* in *MIR156* seedlings were obviously lower than those in WT plants (**Figure 1C**), demonstrating again that the miR156-targeted SPLs enhance ABA responses.

SPL9 Directly Activates the Expression of ABA-Responsive Genes

The above findings that SPLs enhance ABA responses promoted us to study whether SPLs directly bind to the promoters of ABA-responsive genes. As plant-specific transcription factors, SPLs predominantly bind to the common SBP-binding

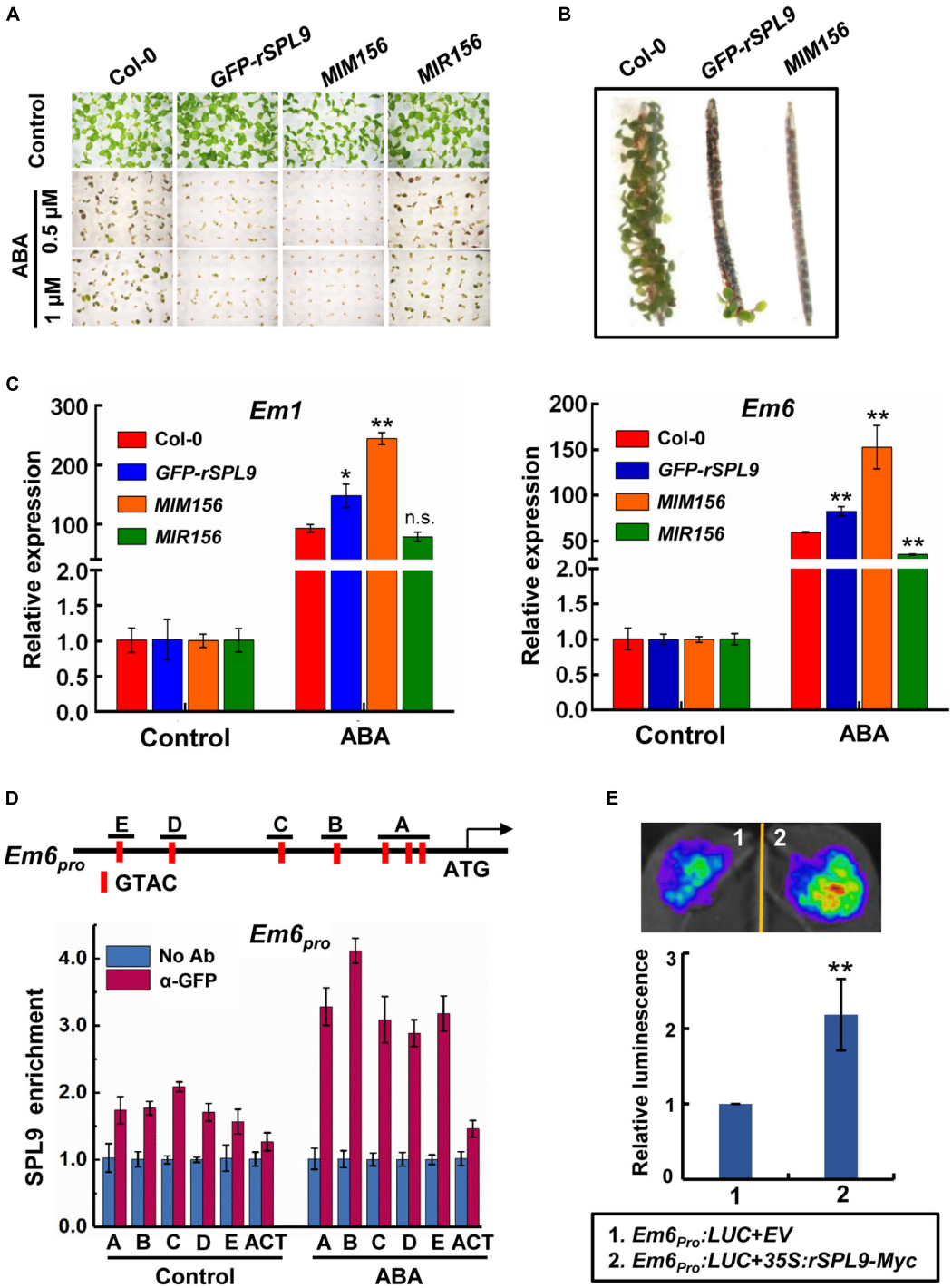

FIGURE 1 | The miR156-targeted SPLs enhance ABA responses. **(A)** Germination phenotype of the indicated seedlings grown on medium containing 0, 0.5, or 1 μM ABA for 7 days. **(B)** Pre-Harvest Sprouting phenotype of the indicated genotypes in fresh mature siliques. **(C)** qRT-PCR assays showing the expression patterns of ABA-responsive genes in 4-day-old seedlings of indicated genotypes with 10 μM ABA treatment (4 h ABA treatment for *Em1*; 2 h ABA treatment for *Em6*). The expression levels in untreated seedlings (Control) for each genotype were set to one. Data are means ± SD ($n = 3$). Asterisks indicate significant differences between the Col-0 and transgenic seedlings. $^{*}P < 0.05$, $^{**}P < 0.01$, n.s. indicates no significant difference (Student *t*-test). **(D)** Chromatin immunoprecipitation (ChIP)-qPCR assays showing the enrichment of SPL9 at the *Em6* promoter regions. The vertical red lines in the upper panel indicate the positions of SBP-box binding core motifs. The 6-day-old *GFP-rSPL9* seedlings treated without (Control) or with 50 μM ABA for 2 h were harvested for ChIP assays. Error bars denote ± SD ($n = 3$). *ACT7* was used as a control. **(E)** Transient expression assays illustrating the activation of *Em6* promoter by SPL9. Upper panel shows a representative leaf image, and the column diagram represents relative luminescence intensities ($n = 15$). The mean value in combination one was set to one. $^{**}P < 0.01$, (Student's *t*-test).

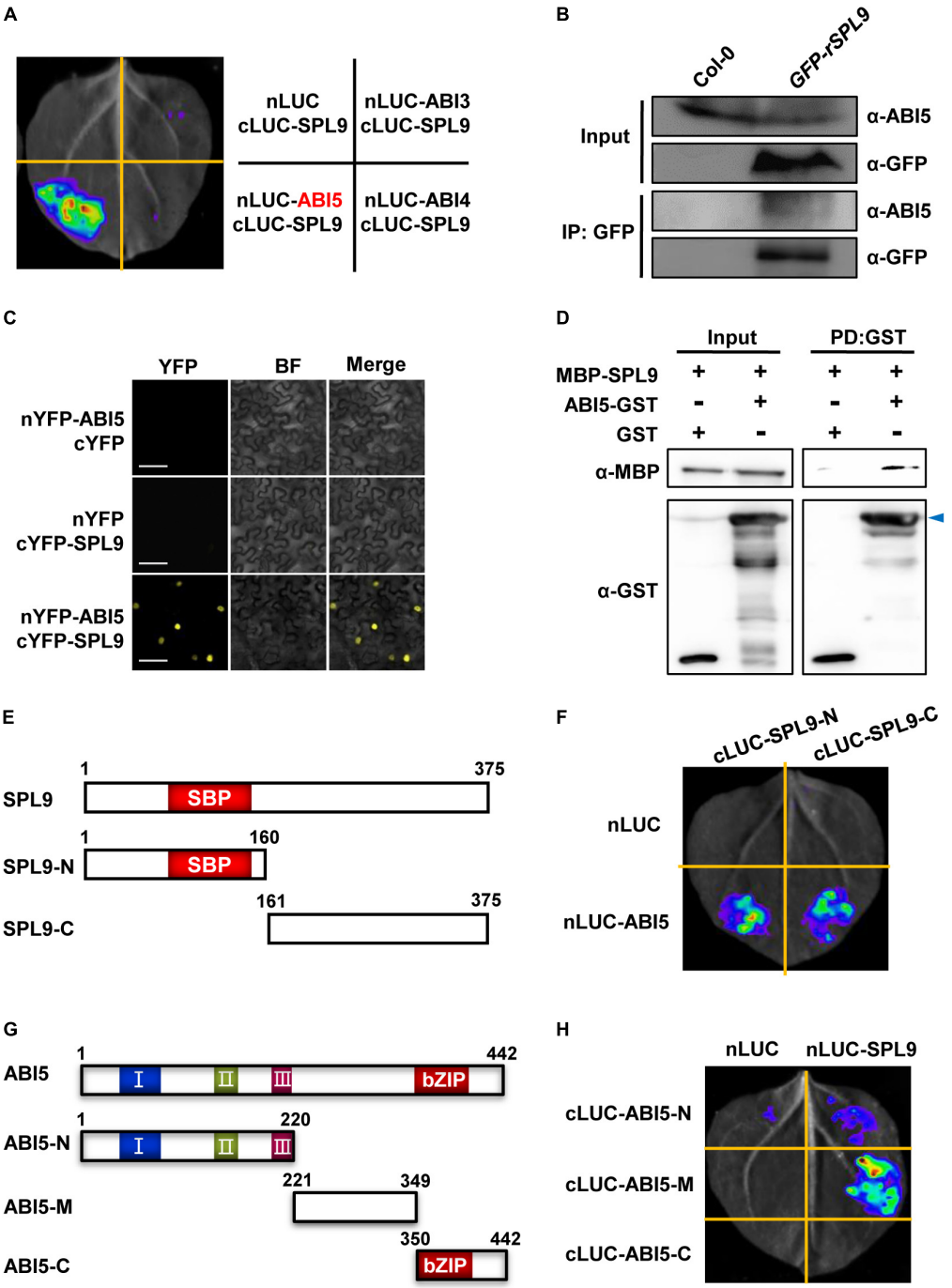

FIGURE 2 | SPL9 physically interacts with ABI5 *in vitro* and *in vivo*. **(A)** Luciferase complementation imaging (LCI) assays showing that SPL9 interacts with ABI5. The cLUC-SPL9 and cLUC-ABI3/ABI4/ABI5 were co-transformed into *Nicotiana benthamiana* leaves, respectively. **(B)** Co-immunoprecipitation (Co-IP) assays showing that SPL9 physically interacts with ABI5 *in vivo*. Total proteins were extracted from the 6-day-old seedlings treated with 50 μM ABA for 2 h. The immunoprecipitates were detected using anti-ABI5 and anti-GFP antibodies, respectively. **(C)** Bimolecular fluorescence complementation (BiFC) assays showing the interaction of SPL9 and ABI5. The constructs indicated were co-transformed into *N. benthamiana* leaves. BF, bright field. Scale bars represent 50 μm. **(D)** Pull-down assays showing that SPL9 directly interacts with ABI5 *in vitro*. Purified MBP-SPL9 proteins could be pulled down by ABI5-GST proteins. MBP was used as a negative control. Arrowhead indicates specific bands. PD, pull down. **(E)** Schematic representation of the full length as well as truncated versions of SPL9 proteins. The N-terminal region of SPL9 contains the SBP domain. **(F)** LCI assays showing the interaction between the truncated SPL9 versions and full-length ABI5. **(G)** Schematic representation of the full length as well as truncated versions of ABI5 proteins. The conserved domains of ABI5 are depicted as I, II, and III (Bensmihen et al., 2002; Lopez-Molina et al., 2003); C-terminal region of ABI5 contains the bZIP domain. **(H)** LCI assays showing the interaction between the truncated ABI5 versions and full-length SPL9.

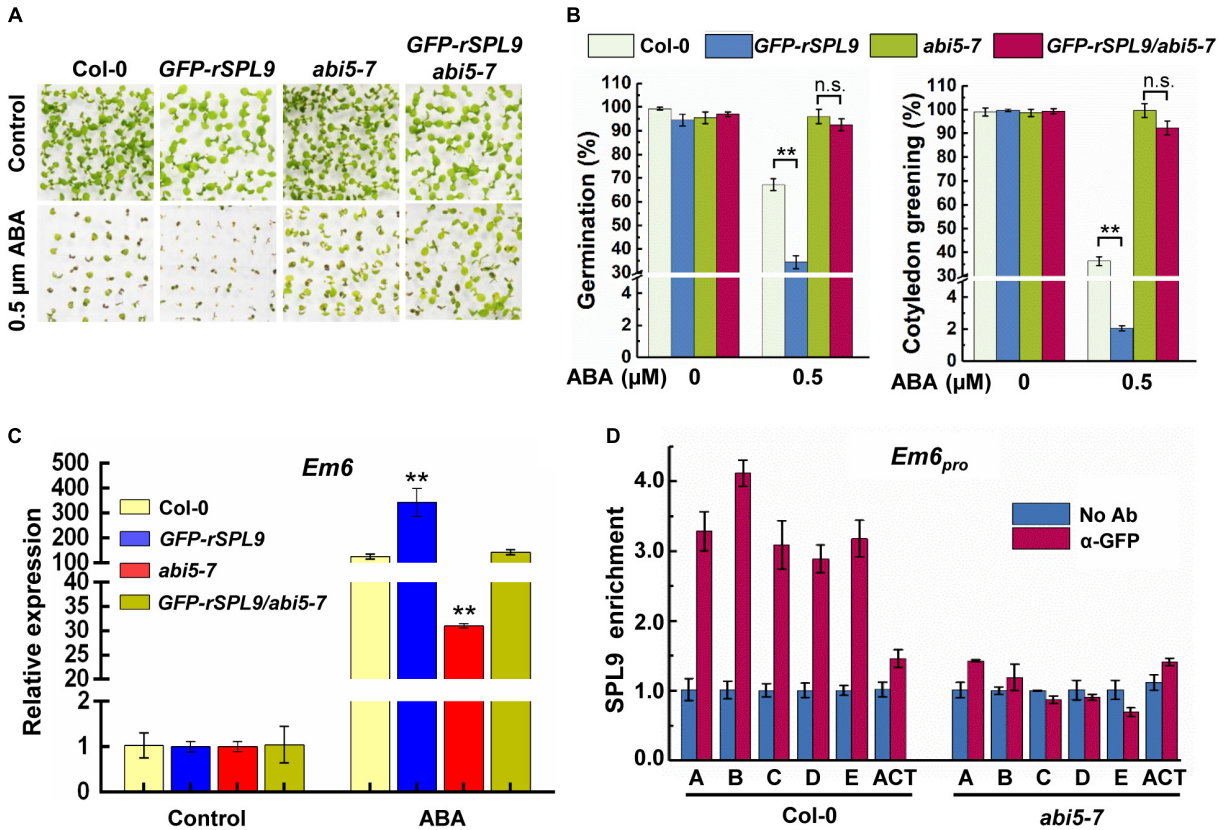

FIGURE 3 | SPL9 promotes ABA responses in an ABI5-dependent manner. **(A)** Germination phenotypes of the indicated seedlings grown on medium without or with 0.5 μM ABA for 7 days. **(B)** Quantification of seed germination and cotyledon greening of indicated genotypes in response to ABA. Seed germination percentage was recorded at 3 days after the end of stratification. Cotyledon-greening percentage was recorded at 5 days after the end of stratification. Data shown are mean ± SD (n = 3). At least 100 seeds per genotype were measured in each replicate. $^{**}P < 0.01$, n.s. indicates no significant difference (Student's t-test). **(C)** qRT-PCR assays showing the expression levels of ABA-responsive gene in the indicated genotypes with ABA treatment. The 4-day-old seedlings were treated without or with 10 μM ABA for 4 h. The expression levels of *Em6* in untreated seedlings (Control) for each genotype were set to one. Data are means ± SD (n = 3). $^{**}P < 0.01$ (Student's t-test). **(D)** ChIP-qPCR assays showing that the ABA-triggered enrichment of SPL9 on the *Em6* promoter is dependent on ABI5. The 6-day-old *GFP-rSPL9* and *GFP-rSPL9/abi5-7* seedlings treated with 50 μM ABA for 2 h were harvested for ChIP assays. Error bars denote ± SD (n = 3). *ACT7* was used as a control.

motifs (such as GTAC) of target genes (Birkenbihl et al., 2005; Liang et al., 2008; Lu et al., 2013). We first scanned the *Em6* (∼1.5-kb) promoter sequence and identified seven putative SBP-binding motifs with positions of −197/−200, −229/−232, −285/−288 (labeled as A), −427/−430 (labeled as B), −626/−629 (labeled as C), −1043/−1046 (labeled as D), and −1220/−1223 (labeled as E), respectively, (**Figure 1D**). Next, we performed chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) assays using the 6-day-old *GFP-rSPL9* seedlings treated without (control) or with 50 μM ABA for 2 h. The results showed that the enrichment of SPL9 at *Em6* promoter was relatively low in the absence of ABA, whereas ABA treatment substantially increased the enrichment of SPL9 at the *Em6* promoter (**Figure 1D**). Similarly, we found two SBP-binding motifs in the *Em1* promoter (800-bp) with positions of −382/−385 (labeled as A) and −675/−678 (labeled as B) (**Supplementary Figure 2A**). The ChIP-qPCR assays showed that SPL9 was also deposited in the *Em1* promoter, especially when treated with ABA (**Supplementary Figure 2A**), implying that ABA can stimulate the deposition of SPL9 to the promoters of ABA-responsive genes.

We further performed transient transcriptional activation assays in *Nicotiana benthamiana* to determine the effect of SPL9 on the transcription of target genes. The *Agrobacterium* strains harboring different constructs, including the $Em6_{pro}$*:LUC* reporter and the effector *35S:rSPL9-Myc*, were co-infiltrated into *N. benthamiana* leaves. The results showed that transient expression of SPL9 could intensely elevate the expression of $Em6_{pro}$*:LUC* reporter (**Figure 1E**). Similarly, the luminescence intensities of $Em1_{pro}$*:LUC* were significantly enhanced when co-expressing with *35S:rSPL9-Myc* (**Supplementary Figure 2B**). These results further suggest that SPL9 could directly activate the transcription of ABA-responsive genes.

SPLs Physically Interact With ABI5

Since SPL9 can directly activate the transcription of ABA-responsive genes, we wondered whether SPL9 interacts with the

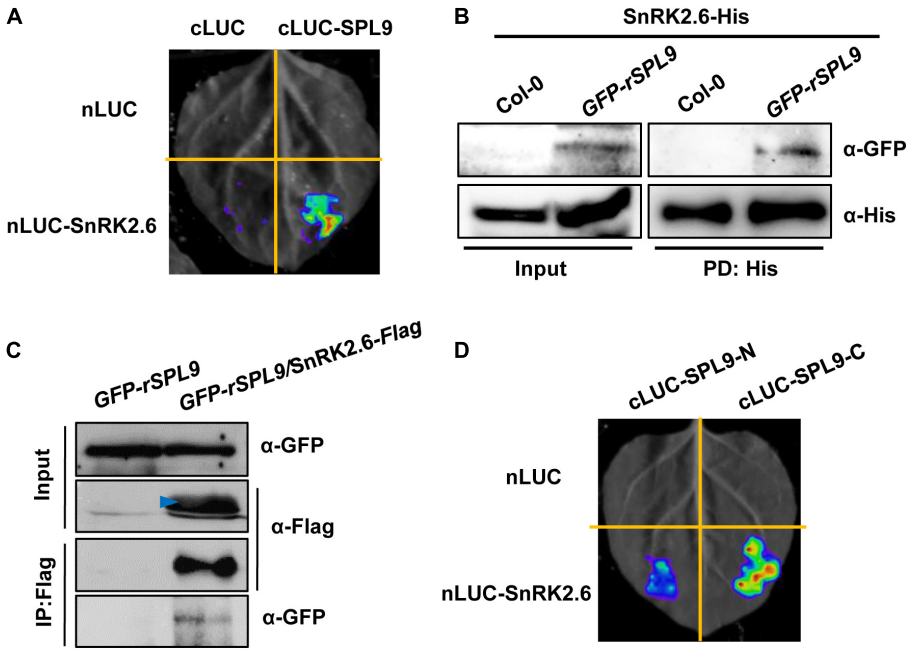

FIGURE 4 | SnRK2.6 directly interacts with SPL9. **(A)** LCI assays showing the interaction between SnRK2.6 and SPL9 in *N. benthamiana* leaves. Empty vectors were used as negative controls. **(B)** Semi-*in vivo* pull-down assays showing the interaction of SnRK2.6 and SPL9. Anti-GFP and anti-His antibodies were used for immunoblotting assays. PD, pull down. **(C)** Co-IP assays showing that SnRK2.6 physically interacts with SPL9 *in vivo*. Total proteins were analyzed by immunoblotting with anti-Flag and anti-GFP antibodies. Arrowhead indicates specific band. **(D)** LCI assays showing the interaction between the truncated SPL9 versions and full-length SnRK2.6.

master transcription factors of ABA signaling, such as ABI3, ABI4 and ABI5. To this end, we performed firefly luciferase complementation imaging (LCI) assays in *N. benthamiana* leaves. ABI3, ABI4, and ABI5 were fused with nLUC to produce nLUC-ABI3/ABI4/ABI5, respectively; meanwhile, SPL9 was fused with cLUC to generate cLUC-SPL9. LCI assays showed that strong luminescence signals were observed in the co-expressed samples of nLUC-ABI5 and cLUC-SPL9, but not in the samples of nLUC-ABI3/cLUC-SPL9, and nLUC-ABI4/cLUC-SPL9 (**Figure 2A**), indicating that SPL9 specifically interacts with ABI5. Furthermore, we conducted co–immunoprecipitation (Co-IP) assays using Col-0 and *GFP-rSPL9* seedlings with ABI5 antibody to confirm the interaction of SPL9 and ABI5. The results showed that ABI5 proteins were co-immunoprecipitated by SPL9 in *GFP-rSPL9* seedlings (**Figure 2B**), suggesting that SPL9 physically interacts with ABI5 *in vivo*. To further confirm the physical interaction between SPL9 and ABI5, we performed bimolecular fluorescence complementation (BiFC) assays in *N. benthamiana* leaves. SPL9 was fused with the C-terminal part of yellow fluorescent protein (cYFP), and ABI5 was fused with the N-terminal part of YFP (nYFP) to generate cYFP-SPL9 and nYFP-ABI5, respectively. The results illustrated that co-expression of cYFP-SPL9 and nYFP-ABI5 produced strong YFP fluorescence in the nucleus, whereas no YFP signal was observed in negative controls (**Figure 2C**). Finally, the pull down assays revealed that GST-ABI5 fusion proteins could retain MBP-SPL9, whereas GST alone could not (**Figure 2D**), suggesting that SPL9 could directly interact with ABI5 *in vitro*. As expected, different approaches demonstrated that SPL3 also interacts with ABI5 (**Supplementary Figures 3A–C**).

Mapping of the Regions Required for the Interaction Between SPL9 and ABI5

To determine which regions of SPL9 are responsible for the interaction with ABI5, we performed LCI assays in *N. benthamiana*. SPL9 was divided into two truncated parts (N: amino-terminal, containing the intact SBP-box domain; C: carboxyl-terminal), according to the position of the highly conserved SBP domain (**Figure 2E**). The results showed that both the N and C termini of SPL9 interact with ABI5 (**Figure 2F**).

Next, to map which region of ABI5 is responsible for the interaction with SPL9, we generated different ABI5 derivatives, including ABI5-N (1-220 aa), ABI5-M (221-349 aa), and ABI5-C (350-442 aa) (**Figure 2G**), based on the highly conserved domains contained in ABI5 (Bensmihen et al., 2002; Lopez-Molina et al., 2003). The results showed that the middle region of ABI5 mediates its interaction with SPL9 (**Figure 2H**).

ABI5 Is Required for the Function of SPL9 in Activating ABA Responses

To evaluate the functional relationship between SPL9 and ABI5 in regulating ABA responses, we generated the *GFP-rSPL9/abi5-7* plants via genetic crossing and examined their seed germination and cotyledon greening phenotypes in response to

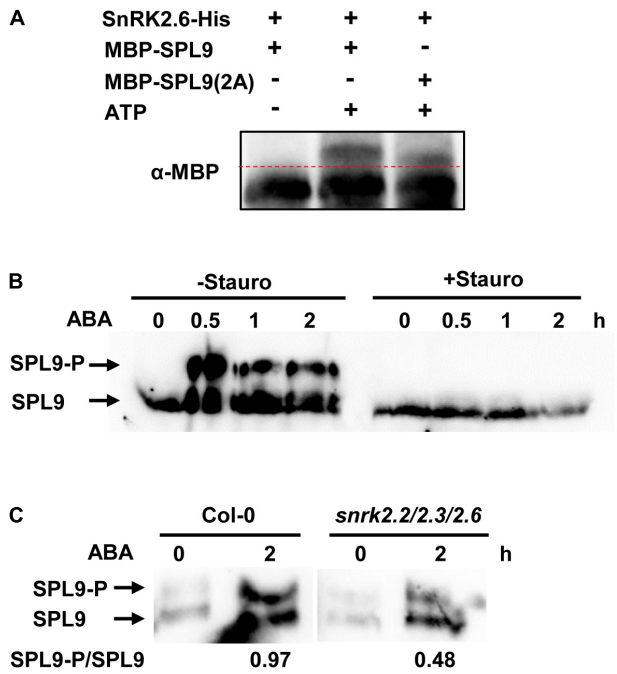

FIGURE 5 | SnRK2.6 phosphorylates SPL9 *in vitro* and *in vivo*. **(A)** *In vitro* kinase assays for the SPL9 protein by SnRK2.6 using phos-tag gel. The mutated SPL9^{S203A,S281A} protein is abbreviated as SPL9(2A). Proteins were detected by immunoblotting with anti-MBP antibody. **(B)** *In vivo* phosphorylation assays for SPL9 using phos-tag gel. The 6-day-old *GFP-rSPL9* seedlings were treated with 50 μM ABA and without or with 5 μM Staurosporine for indicated times. Proteins were analyzed by immunoblotting with anti-GFP antibody. Stauro, Staurosporine. **(C)** *In vivo* phosphorylation assays for SPL9 in indicated genotypes using phos-tag gel. The 6-day-old *GFP-rSPL9* and *GFP-rSPL9/snrk2.2/2.3/2.6* seedlings were treated without or with 50 μM ABA for 2 h. Proteins were analyzed by immunoblotting with anti-GFP antibody. The arrows in **(B)** and **(C)** indicate the phosphorylated or unphosphorylated SPL9.

ABA treatment. Consistent with the above results (**Figure 1A** and **Supplementary Figure 1**), the *GFP-rSPL9* seedlings displayed an ABA-hypersensitive phenotype, whereas the *GFP-rSPL9/abi5-7* and *abi5-7* seedlings displayed decreased sensitivities to ABA treatment compared with Col-0 in terms of seed germination and cotyledon greening (**Figures 3A,B**). This genetic evidence supports the notion that SPL9 enhances ABA responses in an ABI5-dependent manner.

We subsequently determined whether the SPL9-mediated up-regulation of ABA-responsive gene expression is also dependent on ABI5. As expected, the ABA-induced expression of *Em1* and *Em6* in the *GFP-rSPL9* seedlings was significantly increased compared with Col-0, whereas their expression was markedly decreased in the *abi5-7* mutant (**Figure 3C** and **Supplementary Figure 4A**). Intriguingly, the SPL9-enhanced expression of *Em1* and *Em6* was completely suppressed by the *abi5-7* mutation (**Figure 3C** and **Supplementary Figure 4A**). These results promote us to conclude that SPL9 activates the expression of ABA-responsive genes in an ABI5-dependent manner.

The Enrichment of SPL9 at the ABA-Responsive Genes Is Dependent on ABI5

Since ABI5 is required for the function of SPL9 in activating ABA responses, we wondered whether the enrichment of SPL9 at the promoters of ABA-responsive genes is also dependent on ABI5. To this end, we performed ChIP-qPCR assays using the 6-day-old *GFP-rSPL9* and *GFP-rSPL9/abi5-7* seedlings treated with 50 μM ABA for 2 h. Interestingly, the results showed that the ABA-triggered enrichment of SPL9 at the promoters of *Em1* and *Em6* was reduced in the *abi5-7* mutant compared with that in the wild type (**Figure 3D** and **Supplementary Figure 4B**). Notably, the GFP-SPL9 protein levels did not show detectable difference between the wild type and *abi5-7* mutant with or without ABA treatment (**Supplementary Figure 5**). Therefore, we propose that ABI5 facilitates the ABA-triggered recruitment of SPL9 into the chromatin regions of ABA-responsive genes.

SnRK2s Interact With SPLs

Considering the facts that SnRK2s can interact with and phosphorylate ABI5, and SPLs also interact with ABI5, we were curious whether SnRK2s interact with and phosphorylate SPLs. To this end, LCI assays were performed in *N. benthamiana* leaves. As shown in **Figure 4A** and **Supplementary Figure 6**, strong LUC activity was exclusively observed in the co-expressed samples of nLUC-SnRK2s and cLUC-SPL9, indicating that SnRK2s could physically interact with SPL9. We next conducted the semi-*in vitro* pull down assays using the *GFP-rSPL9* seedlings and SnRK2.6-His proteins. The results showed that the GFP-SPL9 fusion proteins were pulled down by SnRK2.6-His proteins (**Figure 4B**). Furthermore, we generated the *GFP-rSPL9/SnRK2.6-Flag* double transgenic plants for Co-IP assays. As shown in **Figure 4C**, the GFP-SPL9 fusion proteins were immunoprecipitated by SnRK2.6-Flag, suggesting that SnRK2.6 interacts with SPL9 *in vivo*. Taken together, these results demonstrate that SnRK2.6 directly interacts with SPL9 *in vitro* and *in vivo*. Meanwhile, SnRK2s could also interact with SPL3 (**Supplementary Figure 7**).

Next, to map which region of SPL9 is responsible for its interaction with SnRK2.6, the full-length SPL9 protein was divided into two parts as described above (**Figure 2E**). The LCI assays in *N. benthamiana* leaves showed that the C terminus of SPL9 predominately mediates the interaction with SnRK2.6 (**Figure 4D**).

SnRK2.6 Phosphorylates SPLs

Since the protein kinase SnRK2.6 interacts with SPL9, we would like to test whether SPL9 is a substrate of SnRK2.6. It has been reported that SnRK2s usually phosphorylate the Ser/Thr residues in the RXXS/T motifs of their substrates (Kobayashi et al., 2005). In this scenario, we first searched the RXXS/T motifs in the SPL9 protein sequence. We found that SPL9 contains two conserved RXXS motifs with putative phosphorylation sites Ser203 and Ser281 (**Supplementary Figure 8**). The *in vitro* phosphorylation assays with the Phos-tag

approach showed that SPL9 could be evidently phosphorylated by SnRK2.6 (**Figure 5A**). Further, we substituted the two putative SnRK2.6 phosphorylation sites of SPL9 with Ala (non-phosphorylated form) to generate the $SPL9^{S203A,S281A}$ mutant form [SPL9(2A)] for *in vitro* phosphorylation assays. As shown in **Figure 5A**, the phosphorylation band of SPL9(2A) was weaker and migrated faster compared with that of the SPL9 protein, indicating that the Ser203 and Ser281 residues are two major SnRK2.6 phosphorylation sites of SPL9. Meanwhile, SnRK2.6 could also phosphorylate SPL3 *in vitro* (**Supplementary Figure 9**).

Furthermore, we wondered whether ABA regulates the SnRK2s-mediated phosphorylation of SPL9 protein. To this end, we used the 6-day-old *GFP-rSPL9* seedlings treated with 50 μM ABA for different time points. Phos-tag gel assays showed that the phosphorylated SPL9 proteins obviously accumulated from 0.5 h after ABA treatment, suggesting that ABA treatment promotes the phosphorylation of SPL9 *in vivo* (**Figure 5B**). Significantly, the ABA-triggered accumulation of phosphorylated SPL9 proteins was almost abolished by the treatment of staurosporine, a general Ser/Thr-kinase inhibitor (**Figure 5B**). To further verify whether the ABA-induced phosphorylation of SPL9 is dependent on the SnRK2s protein kinases, we generated the *GFP-rSPL9/snrk2.2/2.3/2.6* plants by genetic crossing. Phos-tag gel assays showed that the ABA-induced phosphorylation band of SPL9 proteins in the *snrk2.2/2.3/2.6* triple mutant background was much weaker than that in the Col-0 background (**Figure 5C**). These observations suggest that SnRK2s are required for the ABA-induced phosphorylation of SPL9.

SnRK2-Mediated Phosphorylation Is Required for the Function of SPL9 in Enhancing ABA Responses

To elucidate the biological significance of SPL9 phosphorylation by SnRK2s in regulating ABA responses, we generated the $SPL9_{pro}$*:GFP-rSPL9* and $SPL9_{pro}$*:GFP-rSPL9(2A)* transgenic plants. We chose the $SPL9_{pro}$*:GFP-rSPL9* and $SPL9_{pro}$*:GFP-rSPL9(2A)* transgenic lines with similar SPL9 expression levels for further phenotypic analyses (**Figures 6A,B**). As expected, the induction of ABA-responsive genes by ABA in the $SPL9_{pro}$*:GFP-rSPL9* seedlings was higher than that in the wild type (**Figure 6C** and **Supplementary Figure 10A**). Significantly, we found that the ABA induction of ABA-responsive genes in the $SPL9_{pro}$*:GFP-rSPL9(2A)* seedlings was lower than that in the $SPL9_{pro}$*:GFP-rSPL9* seedlings (**Figure 6C** and **Supplementary Figure 10A**), suggesting that the phosphorylation is critical for the function of SPL9 in enhancing ABA responses. To further determine whether SnRK2s is required for the function of SPL9 in enhancing ABA responses, we examined the ABA-induced expression levels of *Em1* and *Em6* in the Col-0, *GFP-rSPL9*, *snrk2.2/2.3/2.6* and *GFP-rSPL9/snrk2.2/2.3/2.6* seedlings. Our results showed that the SPL9-enhanced expression of ABA-responsive genes was abolished in the *snrk2.2/2.3/2.6* triple mutants compared to the wild type (**Figure 6D** and **Supplementary Figure 10B**). The above observations demonstrate that SnRK2s-mediated phosphorylation is required for the activity of SPL9 in enhancing ABA responses. In addition, our results showed the SnRK2s-mediated phosphorylation did not affect the subcellular localization of SPL9 protein in plant cells (**Supplementary Figure 11**).

DISCUSSION

Accumulating evidences have shown that the miR156-SPL regulatory module is highly conserved among different land plant species, and plays important roles in regulating diverse plant developmental processes (Wang and Wang, 2015). Nevertheless, its roles in the ABA signaling remain largely unknown. In this study, we uncover a new biological role of the miR156-SPLs module in regulating ABA response and elucidate the underlying mechanism.

SPLs Activate ABA Signaling in an ABI5-Dependent Manner

The miR156-targeted SPLs have been shown to regulate plant hormone signaling through interacting with several transcription regulators. For example, SPL9 interacts with ARR2, a transcriptional activator of cytokinin signaling, to repress cytokinin response and shoot regeneration (Zhang et al., 2015); SPL9 also interacts with RGA, a transcription repressor of gibberellin (GA) signaling, to regulate flowering time (Yu et al., 2012). However, the role of SPL9 in ABA signaling remains unknown.

In this study, we provide several lines of evidence to demonstrate that the miR156-targeted SPLs facilitate ABA signaling through the interaction with ABI5, a master transcription factor in ABA signaling. First, the miR156-targeted SPLs positively regulate ABA responses (**Figures 1A,C** and **Supplementary Figure 1**). Second, ABA treatment facilitates the recruitment of SPL9 to the promoters of ABA-responsive genes (**Figure 1D** and **Supplementary Figure 2A**). Third, SPLs physically interacts with ABI5 (**Figure 2**). Fourth, genetic analyses reveal that ABI5 is functionally required for SPL9 in activating ABA responses (**Figures 3A–C** and **Supplementary Figure 4A**). Fifth, the ABA-induced enrichment of SPL9 at the promoters of ABA-responsive genes is largely dependent on ABI5 (**Figure 3D** and **Supplementary Figure 4B**). The above-described action mode of SPL9 suppose that SPL9 might function as a cofactor of ABI5 to promote ABA responses. Thus, it is conceivable that the SPLs-mediated enhancement of ABA responses might offer an advantageous strategy for plants to adapt stressful conditions.

SnRK2s Phosphorylate and Activate SPLs During ABA Responses

The SnRK2s family protein kinases act through activation of the transcriptional activity of ABI5 by phosphorylation to promote ABA responses (Nakashima et al., 2009). In this study, we showed that SnRK2s physically interact with and phosphorylate SPLs (**Figures 4**, **5** and **Supplementary Figures 6–9**). We further focused on the biological relevance

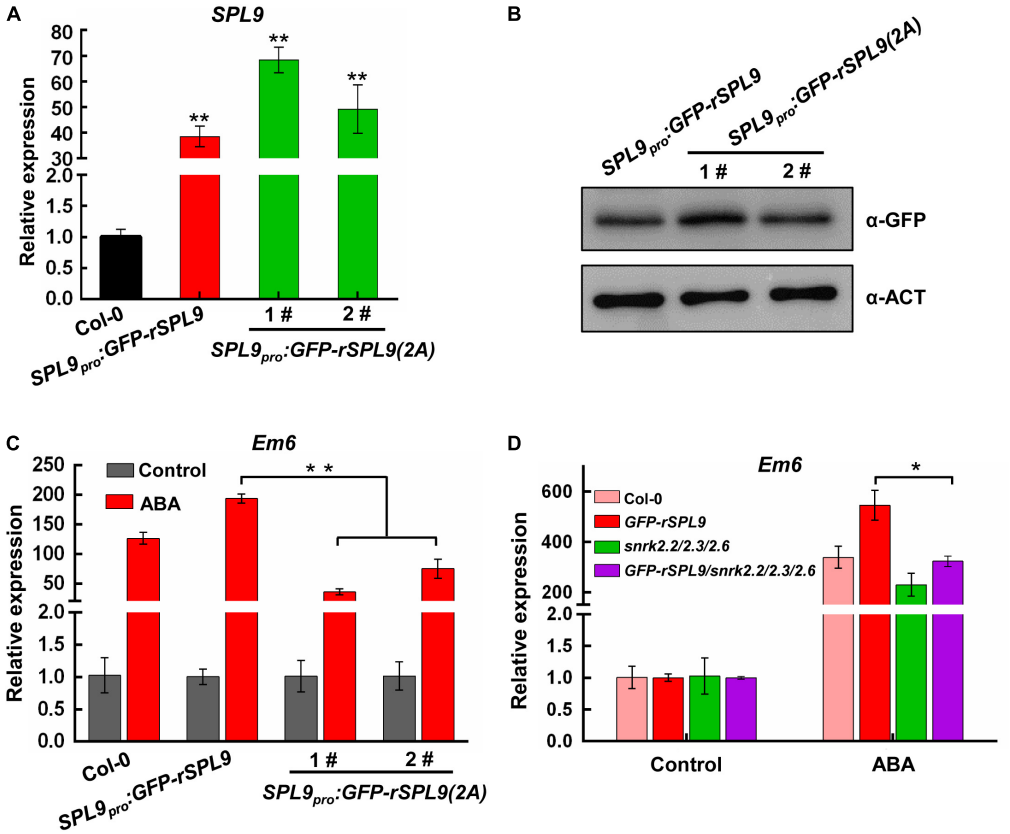

FIGURE 6 | The phosphorylation by SnRK2s is required for SPL9 in enhancing ABA responses. **(A)** qRT-PCR assays showing the expression levels of SPL9 in the indicated seedlings. Values are means ± SD (n = 3). $^{**}P < 0.01$ by the Student's t-test. **(B)** Immunoblotting assays showing the protein levels of SPL9 in the indicated seedlings. Anti-GFP antibody was used. Actin was used as a loading control. **(C)** qRT-PCR assays showing that phosphorylation of SPL9 by SnRK2s is required for the activation of *Em6* expression. The 6-day-old seedlings were treated without or with 50 μM ABA for 4 h. The expression levels of *Em6* in control samples were set to 1 for each genotype. Data are means ± SD (n = 3). $^{**}P < 0.01$ (Student's t-test). **(D)** qRT-PCR assays showing that the SPL9-mediated ABA induction of *Em6* is dependent on SnRK2s. The 4-day-old seedlings were treated without or with 10 μM ABA for 4 h. The expression levels of *Em6* in control samples were set to one for each genotype. Data are means ± SD (n = 3). $^{**}P < 0.01$ (Student's t-test).

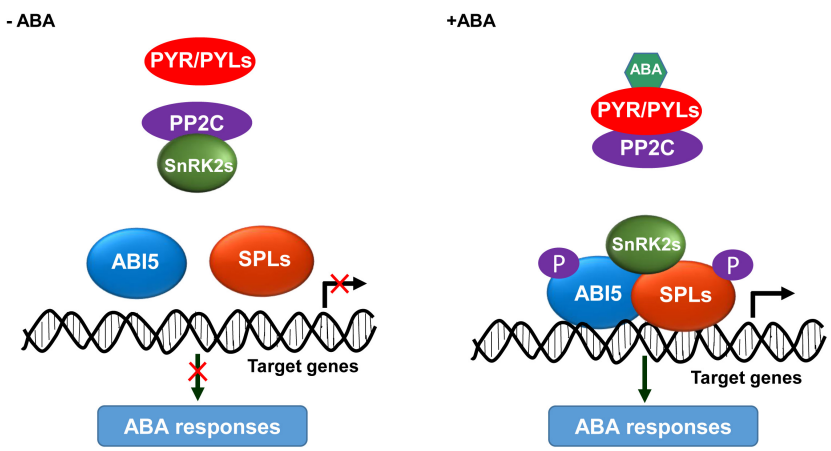

FIGURE 7 | A proposed working model for the SnRK2s-SPLs-ABI5 module in activating ABA responses. In the absence of ABA, PP2C interacts with and dephosphorylates SnRK2s; consequently, ABI5 and SPLs are inactive and unable to activate the ABA responses. In the presence of ABA, its receptors PYR/PYLs interact with PP2C to release the inhibition on SnRK2s activity; thereby, SnRK2s interact with and phosphorylate ABI5 and SPLs, leading to their enrichment at the promoter of target genes to activate ABA responses.

of the phosphorylation of SPL9 by SnRK2s in regulating ABA responses. We found that the expression levels of ABA-responsive genes in the *SPL9*$_{pro}$*:GFP-rSPL9(2A)* seedlings was lower than that in the *SPL9*$_{pro}$*:GFP-rSPL9* seedlings under ABA treatment (**Figure 6C** and **Supplementary Figure 10A**). Notably, both the protein levels and subcellular localization of SPL9 seem not to be affected by the SnRK2s-medidated phosphorylation (**Figure 6B** and **Supplementary Figure 11**). Moreover, genetic analyses showed that the SPL9-mediated ABA induction of ABA-responsive genes was abolished in the absence of *SnRK2.2/2.3/2.6* (**Figure 6D** and **Supplementary Figure 10B**). Taken together, we conclude that the phosphorylation by SnRK2s is essential for SPLs in promoting ABA responses.

The SnRK2s-SPLs-ABI5 Module Is Critical for ABA Signaling

Based on our findings and previous studies (Fujii et al., 2009; Chen et al., 2020), we propose a working model for the mechanism of SnRK2s-SPLs-ABI5 module in activating ABA responses. In the absence of ABA, PP2C dephosphorylates and inactivates SnRK2s; consequently, SPLs and ABI5 are inactive and unable to activate the downstream genes required for ABA responses (**Figure 7**). In the presence of ABA, its receptors PYR/PYLs interact with PP2C to release the inhibition on SnRK2s activity; thereby, the ABA-activated SnRK2s interact with and phosphorylate SPLs and ABI5, leading to their enrichments on the promoter of target genes to activate ABA responses (**Figure 7**). In summary, we discovered that the SnRK2s-SPLs-ABI5 regulatory module represents a signaling hub mediating the enhancement of ABA signaling for plants to adapt to stressful conditions.

Phosphorylation of SPLs in Response to Different Stimuli

Previous studies reported that the *Ideal Plant Architecture 1/Wealthy Farmer's Panicle* (*IPA1/WFP*) gene, encoding an OsSPL14 transcription factor in rice, plays an important role in regulating plant architecture (Jiao et al., 2010; Miura et al., 2010). In addition, the fungus *Magnaporthe oryzae* infection can induce the phosphorylation of OsSPL14, consequently alter its DNA binding specificity (Wang et al., 2018). Unfortunately, the specific protein kinase responsible for the phosphorylation of OsSPL14 in response to *M. oryzae* infection remains to be identified. Significantly, we here found that ABA treatment can induce the phosphorylation of SPL9 by SnRK2s to amplify ABA responses in *Arabidopsis*. Notably, the ABA-induced phosphorylation of SPL9 was reduced in the *snrk2.2/2.3/2.6* triple mutants rather than completely abolished as shown in the wild type seedlings treated with a general Ser/Thr-kinase inhibitor staurosporine, indicating that there might be other protein kinases could phosphorylate SPLs *in vivo*. Thus, we propose that the plant-specific transcription factors SPLs may be phosphorylated and functionally modulated by different protein kinases in response to endogenous cues and external challenges.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

JS designed the research. HD, SY, and YJ performed the experiments. JS, HD, SY, RY, YZhang, YZhou, and YZhu analyzed the data. JS, HD, and SY wrote the manuscript. JS and YZhu revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Central Public-interest Scientific Institute Basic Research Found (grant no. S2021ZD02), the Agricultural Science and Technology Innovation Program of CAAS and the National Natural Science Foundation of China (grant no. 31971880).

ACKNOWLEDGMENTS

We thank Detlef Weigel and Shuhua Yang for sharing research materials.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.708573/full#supplementary-material

Supplementary Figure 1 | The miR156-targeted SPLs positively regulate ABA responses during seed germination and cotyledon greening.

Supplementary Figure 2 | SPL9 directly activates the transcriptional expression of *Em1*.

Supplementary Figure 3 | SPL3 directly interacts with ABI5.

Supplementary Figure 4 | SPL9 activates the expression of *Em1* in an ABI5-dependent manner.

Supplementary Figure 5 | The GFP-SPL9 protein levels were comparable between the *GFP-rSPL9* and *GFP-rSPL9/abi5-7* seedlings.

Supplementary Figure 6 | SnRK2s physically interact with SPL9.

Supplementary Figure 7 | SnRK2s physically interact with SPL3.

Supplementary Figure 8 | The putative SnRK2.6 phosphorylation sites in SPL9 protein.

Supplementary Figure 9 | *In vitro* kinase assay showing that SnRK2.6 could phosphorylate SPL3.

Supplementary Figure 10 | Phosphorylation by SnRK2s is required for SPL9-activated expression of *Em1*.

Supplementary Figure 11 | SnRK2s-mediated phosphorylation did not affect the nuclear localization of SPL9.

Supplementary Table 1 | Primers used for DNA constructs in this study.

Supplementary Table 2 | Constructs used in this study.

Supplementary Table 3 | Primers used for qRT-PCR.

Supplementary Table 4 | Primers used for ChIP-qPCR assays.

REFERENCES

Bensmihen, S., Rippa, S., Lambert, G., Jublot, D., Pautot, V., Granier, F., et al. (2002). The homologous ABI5 and EEL transcription factors function antagonistically to fine-tune gene expression during late embryogenesis. *Plant Cell* 14, 1391–1403. doi: 10.1105/tpc.000869

Birkenbihl, R. P., Jach, G., Saedler, H., and Huijser, P. (2005). Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. *J. Mol. Biol.* 352, 585–596. doi: 10.1016/j.jmb.2005.07.013

Bu, Q., Li, H., Zhao, Q., Jiang, H., Zhai, Q., Zhang, J., et al. (2009). The *Arabidopsis* RING finger E3 ligase RHA2a is a novel positive regulator of abscisic acid signaling during seed germination and early seedling development. *Plant Physiol.* 150, 463–481. doi: 10.1104/pp.109.135269

Cardon, G., Hohmann, S., Klein, J., Nettesheim, K., Saedler, H., and Huijser, P. (1999). Molecular characterisation of the *Arabidopsis* SBP-box genes. *Gene* 237, 91–104. doi: 10.1016/s0378-1119(99)00308-x

Chen, H., Zou, Y., Shang, Y., Lin, H., Wang, Y., Cai, R., et al. (2008). Firefly luciferase complementation imaging assay for protein-protein interactions in plants. *Plant Physiol.* 146, 368–376.

Chen, K., Li, G. J., Bressan, R. A., Song, C. P., Zhu, J. K., and Zhao, Y. (2020). Abscisic acid dynamics, signaling, and functions in plants. *J. Integr. Plant Biol.* 62, 25–54. doi: 10.1111/jipb.12899

Chen, R., Jiang, H., Li, L., Zhai, Q., Qi, L., Zhou, W., et al. (2012). The *Arabidopsis* mediator subunit MED25 differentially regulates jasmonate and abscisic acid signaling through interacting with the MYC2 and ABI5 transcription factors. *Plant Cell* 24, 2898–2916. doi: 10.1105/tpc.112.098277

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

Cutler, S. R., Rodriguez, P. L., Finkelstein, R. R., and Abrams, S. R. (2010). Abscisic acid: emergence of a core signaling network. *Annu. Rev. Plant Biol.* 61, 651–679. doi: 10.1146/annurev-arplant-042809-112122

Ding, Y., Li, H., Zhang, X., Xie, Q., Gong, Z., and Yang, S. (2015). OST1 kinase modulates freezing tolerance by enhancing ICE1 stability in *Arabidopsis*. *Dev. Cell* 32, 278–289. doi: 10.1016/j.devcel.2014.12.023

Dong, H., Liu, J., He, G., Liu, P., and Sun, J. (2020). Photoexcited phytochrome B interacts with brassinazole resistant 1 to repress brassinosteroid signaling in *Arabidopsis*. *J. Integr. Plant Biol.* 62, 652–667. doi: 10.1111/jipb.12822

Finkelstein, R., Lynch, T., Reeves, W., Petitfils, M., and Mostachetti, M. (2011). Accumulation of the transcription factor ABA-insensitive (ABI)4 is tightly regulated post-transcriptionally. *J. Exp. Bot.* 62, 3971–3979. doi: 10.1093/jxb/err093

Finkelstein, R. R., and Lynch, T. J. (2000). The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. *Plant Cell* 12, 599–609. doi: 10.2307/3871072

Finkelstein, R. R., Wang, M. L., Lynch, T. J., Rao, S., and Goodman, H. M. (1998). The *Arabidopsis* abscisic acid response locus ABI4 encodes an APETALA 2 domain protein. *Plant Cell* 10, 1043–1054. doi: 10.2307/3870689

Fujii, H., Chinnusamy, V., Rodrigues, A., Rubio, S., Antoni, R., Park, S. Y., et al. (2009). In vitro reconstitution of an abscisic acid signalling pathway. *Nature* 462, 660–664. doi: 10.1038/nature08599

Fujii, H., Verslues, P. E., and Zhu, J. K. (2007). Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in *Arabidopsis*. *Plant Cell* 19, 485–494. doi: 10.1105/tpc.106.048538

Fujii, H., and Zhu, J. K. (2009). *Arabidopsis* mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8380–8385. doi: 10.1073/pnas.0903144106

Giraudat, J., Hauge, B. M., Valon, C., Smalle, J., Parcy, F., and Goodman, H. M. (1992). Isolation of the *Arabidopsis* ABI3 gene by positional cloning. *Plant Cell* 4, 1251–1261. doi: 10.2307/3869411

Gosti, F., Beaudoin, N., Serizet, C., Webb, A. A. R., Vartanian, N., and Giraudat, J. (1999). ABI1 protein phosphatase 2C is a negative regulator of abscisic acid signaling. *Plant Cell* 11, 1897–1909. doi: 10.2307/3871085

Huang, X., Hou, L., Meng, J., You, H., Li, Z., Gong, Z., et al. (2018). The antagonistic action of abscisic acid and cytokinin signaling mediates drought stress response in *Arabidopsis*. *Mol. Plant* 11, 970–982. doi: 10.1016/j.molp.2018.05.001

Huo, H. Q., Wei, S. H., and Bradford, K. J. (2016). DELAY OF GERMINATION1 (DOG1) regulates both seed dormancy and flowering time through microRNA pathways. *Proc. Natl. Acad. Sci. U.S.A.* 113, E2199–E2206.

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

Kinoshita, E., Kinoshita-Kikuta, E., Takiyama, K., and Koike, T. (2006). Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol. Cell. Proteomics* 5, 749–757. doi: 10.1074/mcp.t500024-mcp200

Kobayashi, Y., Murata, M., Minami, H., Yamamoto, S., Kagaya, Y., Hobo, T., et al. (2005). Abscisic acid-activated SNRK2 protein kinases function in the gene-regulation pathway of ABA signal transduction by phosphorylating ABA response element-binding factors. *Plant J.* 44, 939–949. doi: 10.1111/j.1365-313x.2005.02583.x

Leung, J., Bouvier-Durand, M., Morris, P. C., Guerrier, D., Chefdor, F., and Giraudat, J. (1994). *Arabidopsis* ABA response gene ABI1: features of a calcium-modulated protein phosphatase. *Science* 264, 1448–1452. doi: 10.1126/science.7910981

Leung, J., Merlot, S., and Giraudat, J. (1997). The *Arabidopsis* ABSCISIC ACID-INSENSITIVE2 (ABI2) and ABI1 genes encode homologous protein phosphatases 2C involved in abscisic acid signal transduction. *Plant Cell* 9, 759–771. doi: 10.2307/3870430

Li, H., Jiang, H., Bu, Q., Zhao, Q., Sun, J., Xie, Q., et al. (2011). The *Arabidopsis* RING finger E3 ligase RHA2b acts additively with RHA2a in regulating abscisic acid signaling and drought response. *Plant Physiol.* 156, 550–563. doi: 10.1104/pp.111.176214

Liang, X., Nazarenus, T. J., and Stone, J. M. (2008). Identification of a consensus DNA-binding site for the *Arabidopsis thaliana* SBP domain transcription factor, AtSPL14, and binding kinetics by surface plasmon resonance. *Biochemistry* 47, 3645–3653. doi: 10.1021/bi701431y

Liu, M. M., Shi, Z. Y., Zhang, X. H., Wang, M. X., Zhang, L., Zheng, K. Z., et al. (2019). Inducible overexpression of ideal plant architecture1 improves both yield and disease resistance in rice. *Nat. Plants* 5, 389–400. doi: 10.1038/s41477-019-0383-2

Lopez-Molina, L., Mongrand, S., Kinoshita, N., and Chua, N. H. (2003). AFP is a novel negative regulator of ABA signaling that promotes ABI5 protein degradation. *Genes Dev.* 17, 410–418. doi: 10.1101/gad.1055803

Lu, Q., Tang, X., Tian, G., Wang, F., Liu, K., Nguyen, V., et al. (2010). *Arabidopsis* homolog of the yeast TREX-2 mRNA export complex: components and anchoring nucleoporin. *Plant J.* 61, 259–270. doi: 10.1111/j.1365-313x.2009.04048.x

Lu, Z., Yu, H., Xiong, G., Wang, J., Jiao, Y., Liu, G., et al. (2013). Genome-wide binding analysis of the transcription activator ideal plant architecture1 reveals a complex network regulating rice plant architecture. *Plant Cell* 25, 3743–3759. doi: 10.1105/tpc.113.113639

Ma, Y., Szostkiewicz, I., Korte, A., Moes, D., Yang, Y., Christmann, A., et al. (2009). Regulators of PP2C phosphatase activity function as abscisic acid sensors. *Science* 324, 1064–1068.

Melcher, K., Ng, L. M., Zhou, X. E., Soon, F. F., Xu, Y., Suino-Powell, K. M., et al. (2009). A gate-latch-lock mechanism for hormone signalling by abscisic acid receptors. *Nature* 462, 602–608. doi: 10.1038/nature08613

Miao, C. B., Wang, Z., Zhang, L., Yao, J. J., Hua, K., Liu, X., et al. (2019). The grain yield modulator miR156 regulates seed dormancy through the gibberellin pathway in rice. *Nat. Commu.* 10:3822.

Miura, K., Ikeda, M., Matsubara, A., Song, X. J., Ito, M., Asano, K., et al. (2010). OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat. Genet.* 42, 545–549. doi: 10.1038/ng.592

Miyazono, K., Miyakawa, T., Sawano, Y., Kubota, K., Kang, H. J., Asano, A., et al. (2009). Structural basis of abscisic acid signalling. *Nature* 462, 609–614. doi: 10.1038/nature08583

Mustilli, A. C., Merlot, S., Vavasseur, A., Fenzi, F., and Giraudat, J. (2002). *Arabidopsis* OST1 protein kinase mediates the regulation of stomatal aperture by abscisic acid and acts upstream of reactive oxygen species production. *Plant Cell* 14, 3089–3099. doi: 10.1105/tpc.007906

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., et al. (2007). Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* 104, 34–41. doi: 10.1263/jbb.104.34

Nakashima, K., Fujita, Y., Kanamori, N., Katagiri, T., Umezawa, T., Kidokoro, S., et al. (2009). Three *Arabidopsis* SnRK2 protein kinases, SRK2D/SnRK2.2, SRK2E/SnRK2.6/OST1 and SRK2I/SnRK2.3, involved in ABA signaling are essential for the control of seed development and dormancy. *Plant Cell Physiol.* 50, 1345–1363. doi: 10.1093/pcp/pcp083

Park, S. Y., Fung, P., Nishimura, N., Jensen, D. R., Fujii, H., Zhao, Y., et al. (2009). Abscisic acid inhibits type 2C protein phosphatases via the PYR/PYL family of START proteins. *Science* 324, 1068–1071.

Rodriguez, P. L., Benning, G., and Grill, E. (1998). ABI2, a second protein phosphatase 2C involved in abscisic acid signal transduction in *Arabidopsis*. *FEBS Lett.* 421, 185–190. doi: 10.1016/s0014-5793(97)01558-5

Santiago, J., Dupeux, F., Round, A., Antoni, R., Park, S. Y., Jamin, M., et al. (2009). The abscisic acid receptor PYR1 in complex with abscisic acid. *Nature* 462, 665–668. doi: 10.1038/nature08591

Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., and Weigel, D. (2005). Specific effects of microRNAs on the plant transcriptome. *Dev. Cell* 8, 517–527. doi: 10.1016/j.devcel.2005.01.018

Schwarz, S., Grande, A. V., Bujdoso, N., Saedler, H., and Huijser, P. (2008). The microRNA regulated SBP-box genes SPL9 and SPL15 control shoot maturation in *Arabidopsis*. *Plant Mol. Biol.* 67, 183–195. doi: 10.1007/s11103-008-9310-z

Soon, F. F., Ng, L. M., Zhou, X. E., West, G. M., Kovach, A., Tan, M. H., et al. (2012). Molecular mimicry regulates ABA signaling by SnRK2 kinases and PP2C phosphatases. *Science* 335, 85–88. doi: 10.1126/science.1215106

Sun, J., Qi, L., Li, Y., Chu, J., and Li, C. (2012). PIF4-mediated activation of YUCCA8 expression integrates temperature into the auxin pathway in regulating *Arabidopsis* hypocotyl growth. *PLoS Genet.* 8:e1002594. doi: 10.1371/journal.pgen.1002594

Umezawa, T., Nakashima, K., Miyakawa, T., Kuromori, T., Tanokura, M., Shinozaki, K., et al. (2010). Molecular basis of the core regulatory network in ABA responses: sensing, signaling and transport. *Plant Cell Physiol.* 51, 1821–1839. doi: 10.1093/pcp/pcq156

Umezawa, T., Sugiyama, N., Mizoguchi, M., Hayashi, S., Myouga, F., Yamaguchi-Shinozaki, K., et al. (2009). Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17588–17593. doi: 10.1073/pnas.0907095106

Wang, H., and Wang, H. (2015). The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. *Mol. Plant* 8, 677–688. doi: 10.1016/j.molp.2015.01.008

Wang, J., Zhou, L., Shi, H., Chern, M., Yu, H., Yi, H., et al. (2018). A single transcription factor promotes both yield and immunity in rice. *Science* 361, 1026–1028. doi: 10.1126/science.aat7675

Wang, J. W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell* 138, 738–749. doi: 10.1016/j.cell.2009.06.014

Weiner, J. J., Peterson, F. C., Volkman, B. F., and Cutler, S. R. (2010). Structural and functional insights into core ABA signaling. *Curr. Opin. Plant Biol.* 13, 495–502. doi: 10.1016/j.pbi.2010.09.007

Wu, G., and Poethig, R. S. (2006). Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development* 133, 3539–3547. doi: 10.1242/dev.02521

Xie, Y., Liu, Y., Wang, H., Ma, X., Wang, B., Wu, G., et al. (2017). Phytochrome-interacting factors directly suppress MIR156 expression to enhance shade-avoidance syndrome in *Arabidopsis*. *Nat. Commun.* 8:348.

Xing, S., Salinas, M., Hohmann, S., Berndtgen, R., and Huijser, P. (2010). miR156-targeted and nontargeted SBP-box transcription factors act in concert to secure male fertility in *Arabidopsis*. *Plant Cell* 22, 3935–3950. doi: 10.1105/tpc.110.079343

Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., et al. (2004). A novel zinc-binding motif revealed by solution structures of DNA-binding domains of *Arabidopsis* SBP-family transcription factors. *J. Mol. Biol.* 337, 49–63. doi: 10.1016/j.jmb.2004.01.015

Yang, Z., Wang, X., Gu, S., Hu, Z., Xu, H., and Xu, C. (2008). Comparative study of SBP-box gene family in *Arabidopsis* and rice. *Gene* 407, 1–11. doi: 10.1016/j.gene.2007.02.034

Yu, S., Galvao, V. C., Zhang, Y. C., Horrer, D., Zhang, T. Q., Hao, Y. H., et al. (2012). Gibberellin regulates the *Arabidopsis* floral transition through miR156-targeted SQUAMOSA PROMOTER BINDING-LIKE transcription factors. *Plant Cell* 24, 3320–3332. doi: 10.1105/tpc.112.101014

Zhang, T. Q., Lian, H., Tang, H., Dolezal, K., Zhou, C. M., Yu, S., et al. (2015). An intrinsic microRNA timer regulates progressive decline in shoot regenerative capacity in plants. *Plant Cell* 27, 349–360. doi: 10.1105/tpc.114.135186

Zhu, J. Y., Sun, Y., and Wang, Z. Y. (2012). Genome-wide identification of transcription factor-binding sites in plants using chromatin immunoprecipitation followed by microarray (ChIP-chip) or sequencing (ChIP-seq). *Methods Mol. Biol.* 876, 173–188. doi: 10.1007/978-1-61779-809-2_14

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 in Plant Science

ORIGINAL RESEARCH
published: 11 August 2021
doi: 10.3389/fpls.2021.717649

LFR Physically and Genetically Interacts With SWI/SNF Component SWI3B to Regulate Leaf Blade Development in Arabidopsis

Xiaowei Lin[1,2†], Can Yuan[1†], Bonan Zhu[1], Tingting Yuan[1], Xiaorong Li[1], Shan Yuan[1], Sujuan Cui[1] and Hongtao Zhao[1]**

[1] Hebei Key Laboratory of Molecular and Cellular Biology, Key Laboratory of Molecular and Cellular Biology of Ministry of Education, Hebei Collaboration Innovation Center for Cell Signaling and Environmental Adaptation, College of Life Sciences, Hebei Normal University, Shijiazhuang, China, [2] School of Traditional Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin, China

OPEN ACCESS

Edited by:
Dazhong Dave Zhao,
University of Wisconsin–Milwaukee,
United States

Reviewed by:
Gyung-Tae Kim,
Dong-A University, South Korea
Keqiang Wu,
National Taiwan University, Taiwan
Chenlong Li,
Agriculture and Agri-Food Canada (AAFC), Canada

***Correspondence:**
Sujuan Cui
cuisujuan@hebtu.edu.cn
Hongtao Zhao
zhaohongtao@hebtu.edu.cn

[†] These authors have contributed equally to this work

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

Received: 31 May 2021
Accepted: 14 July 2021
Published: 11 August 2021

Citation:
Lin X, Yuan C, Zhu B, Yuan T, Li X, Yuan S, Cui S and Zhao H (2021) LFR Physically and Genetically Interacts With SWI/SNF Component SWI3B to Regulate Leaf Blade Development in Arabidopsis. Front. Plant Sci. 12:717649. doi: 10.3389/fpls.2021.717649

Leaves start to develop at the peripheral zone of the shoot apical meristem. Thereafter, symmetric and flattened leaf laminae are formed. These events are simultaneously regulated by auxin, transcription factors, and epigenetic regulatory factors. However, the relationships among these factors are not well known. In this study, we conducted protein-protein interaction assays to show that our previously reported Leaf and Flower Related (LFR) physically interacted with SWI3B, a component of the ATP-dependent chromatin remodeling SWI/SNF complex in Arabidopsis. The results of truncated analysis and transgenic complementation showed that the N-terminal domain (25–60 amino acids) of LFR was necessary for its interaction with SWI3B and was crucial for LFR functions in Arabidopsis leaf development. Genetic results showed that the artificial microRNA knockdown lines of *SWI3B* (*SWI3B-amic*) had a similar upward-curling leaf phenotype with that of *LFR* loss-of-function mutants. ChIP-qPCR assay was conducted to show that LFR and SWI3B co-targeted the promoters of *YABBY1/FILAMENTOUS FLOWER* (*YAB1/FIL*) and *IAA carboxyl methyltransferase 1* (*IAMT1*), which were misexpressed in *lfr* and *SWI3B-amic* mutants. In addition, the association between LFR and the *FIL* and *IAMT1* loci was partly hampered by the knockdown of *SWI3B*. These data suggest that LFR interacts with the chromatin-remodeling complex component, SWI3B, and influences the transcriptional expression of the important transcription factor, *FIL*, and the auxin metabolism enzyme, *IAMT1*, in flattened leaf lamina development.

Keywords: LFR, SWI3B, *FIL*, *IAMT1*, SWI/SNF, leaf, Arabidopsis

INTRODUCTION

Leaves are the main sites of photosynthesis, a process that results in the production of food in plants, which are then consumed by animals. Leaf morphology is an important trait that affects the efficiency of photosynthesis and crop yield. Leaves develop from leaf primordia, which are located in the peripheral zone of the shoot apical meristem (SAM). The polarity of leaf primordia along the adaxial-abaxial, proximal-distal, and medio-lateral axes are first established (McConnell and Barton, 1998; Bowman et al., 2002; Du et al., 2018). Cells that are destined to appear on the adaxial side of the leaf are determined by HD-ZIP III and related transcription factors, while those that are destined to appear on the abaxial side of the leaf are established and maintained by

YABBY (YAB) and KANADI (KAN) transcription factors. These adaxial and abaxial cell fate regulators are coordinated by auxin and a transcription factor called ASYMMETRIC LEAVES2 (AS2), which act on flattened leaves during their development (Wu et al., 2008; Jun et al., 2010; Husbands et al., 2015; Manuela and Xu, 2020). However, the epigenetic regulatory mechanisms of these regulators and their effect on leaf development should be elucidated.

In eukaryotes, ATP-dependent chromatin remodeling complexes (CRCs) are a group of crucial epigenetic factors that utilize energy from ATP hydrolysis to influence chromatin or nucleosome conformation and transcriptional gene expression (Vignali et al., 2000; Hargreaves and Crabtree, 2011). As a conserved subfamily of CRCs, the SWITCHING/SUCROSE NON-FERMENTING (SWI/SNF) complex usually contains four conserved core subunits, including Swi2/Snf2 ATPase, Swi3, Snf5, and Swp73/BAF60/CHC. These core subunits are required for the assembly and activity of the SWI/SNF complex (Sudarsanam and Winston, 2000; Yang et al., 2007; Sundaramoorthy and Owen-Hughes, 2020). Several core subunits of the plant Swi2/Snf2 ATPase BRAHMA (BRM)-SWI/SNF complex, such as BRAHMA-interacting proteins 1 (BRIP1) and BRIP2, and bromodomain-containing proteins BRD1, BRD2, and BRD13, have recently been discovered to co-localize and act together with BRM on chromatin to regulate gene expression (Yu et al., 2020, 2021). In the genome of the model plant, *Arabidopsis thaliana*, there are four Swi3 proteins, including SWI3A/3B/3C/3D (Sarnowski et al., 2005). Results of a genetic analysis indicate that these components play essential roles in regulating multiple growth and developmental processes (Sarnowski et al., 2005; Han et al., 2018; Jiang et al., 2019; Yang et al., 2020). SWI3A, SWI3B, and SWI3C proteins interact with one another, whereas SWI3D only interacts with SWI3B (Sarnowski et al., 2005). Additionally, SWI3B interacts with a long non-coding (lnc)RNA-binding protein called INVOLVED IN *DE NOVO* 2 (IDN2) or with histone deacetylase HISTONE DEACETYLASE 6 (HDA6) to maintain non-coding RNA-mediated transcriptional or transposon silencing (Zhu et al., 2013; Yang et al., 2020). Moreover, SWI3C is involved in the regulation of leaf size in Arabidopsis and tomato (Vercruyssen et al., 2014; Zhao et al., 2019). Arabidopsis SWI3C and BRM interact with the transcription factor, TEOSINTE BRANCHED1, CYCLOIDEA, PCF4 (TCP4), to promote cell differentiation in leaves by increasing the transcriptional expression of *ARABIDOPSIS RESPONSE REGULATOR 16* (*ARR16*), an inhibitor of cytokinin response (Efroni et al., 2013). Embryos of the null mutants of *SWI3B* genes exhibited early lethality (Sarnowski et al., 2005), whereas knockdown mutants of *SWI3B* with RNA interference (*SWI3B-RNAi*) resulted in an upward-curling leaf phenotype (Han et al., 2018). The increased transcript level and decreased nucleosome occupation of *IAA carboxyl methyltransferase 1* (*IAMT1*) may explain this defect observed during the development of leaves with *SWI3B-RNAi* (Han et al., 2018). However, the direct targets of SWI3B and its interacting partners in leaf development still need clarification.

The *Leaf and Flower-Related gene* (*LFR*) encodes a nuclear protein with the Armadillo (ARM)-repeat domains (Wang et al., 2009), which are involved in protein–protein interactions (Samuel et al., 2006). Arabidopsis with a loss-of-function mutation in the *LFR* gene exhibit pleiotropic phenotypes during leaf and flower development (Wang et al., 2009, 2012; Lin et al., 2018). LFR has been isolated from tandem affinity-purified protein complexes using SWIP37B (Vercruyssen et al., 2014). It interacts genetically and physically with AS2 to co-repress the transcription expression of *BREVIPEDICELLUS* (*BP*), which influences chromatin configuration during the determination of petiole length, vasculature pattern, and leaf margin development (Lin et al., 2018). However, the interacting partners and downstream targets of LFR during the development of flattened lamina remain largely unknown.

This study aimed to determine the interacting partner of LFR, examine the physical and genetic relationships between LFR and SWI3B during flattened leaf development in Arabidopsis, detect changes in the expression of the *FIL* and *IAMT1* genes in Arabidopsis with single mutant of *lfr* and in those with knockdown mutants of *SW13B*, and investigate the binding peaks of LFR and SW13B in the *FIL* and *IAMT1* promoter regions.

MATERIALS AND METHODS

Plant Growth Conditions

We used *A. thaliana*, the commonly used and well-studied model plant, in this study. All Arabidopsis plants in this study had a Columbia-0 background. The seeds of *lfr-1* and *lfr-2*/+ were previously reported in our laboratory (Wang et al., 2009). *swi3b-2*/+ were previously reported (Sarnowski et al., 2005). Other transgenic plants were obtained in this sturdy by floral infiltration (Clough and Bent, 1998), after successful plasmid constructions described in the next part. The seeds were surface-sterilized with 75% ethanol, stored at 4°C for 3 days, and cultured on Murashige and Skoog (MS) medium containing 1% sucrose (pH 5.7). After 10 days of growth, the seedlings were transplanted into soil and grown in a greenhouse under a 16-h light/8-h dark photoperiod at 22°C.

Plasmid Constructions

For the binary vectors for the transgenic complementation and genetic analysis, the coding sequences of the full or truncated *LFR*, *SWI3B*, and *FIL* were amplified with specific primers (**Supplementary Table 1**) using the plasmid pTR5 (for *LFR*) (Wang et al., 2009) or cDNA (for *SWI3B* and *FIL*) as the template. The amplified fragment was digested using an appropriate restriction endonuclease and inserted into p*CAMBIA1300 35S:3FLAG* to obtain p*35S:LFR* (full length or truncated)-*3FLAG*, *35S:SWI3B-3FLAG*, and *35S:FIL-3FLAG*.

The yeast two-hybrid (Y2H) GAL4 system bait/prey plasmid, which had a coding sequence of full or truncated LFR or SWI3B were separately constructed. Briefly, the coding sequences of the full or truncated *LFR* or *SWI3B* were amplified with specific primers (**Supplementary Table 1**) using the plasmid, pTR5 (Wang et al., 2009), and cDNA as a template. The amplified fragment was digested using an appropriate restriction endonuclease and inserted into prey pGADT7/bait pGBKT7

to obtain pGADT7/pGBKT7-LFR (full length or truncated), pGADT7/pGBKT7-SWI3B (full length or truncated).

In bimolecular fluorescence complementation (BiFC) experiments, full-length CDS of SWI3B with a stop codon was amplified via polymerase chain reaction (PCR) using the Arabidopsis cDNA as a template and cloned into pENTRY/D/SD-TOPO. These genes were then introduced into pxnYFPGW via the LR reaction. The N terminal part of nYFP-AS2 and the C terminal part of CFP-LFR (cCFP-LFR) plasmids were reported in our previous study (Lin et al., 2018). The specific primers used for plasmid construction are listed in **Supplementary Table 1**.

For artificial miRNA construction of *SWI3B*, artificial miRNA site selection, primers, and specific construction procedures were carried out according to the description on the Web of MicroRNA Designer platform (WMD)[1]. The artificial miRNA precursors, *mic1* and *mic2*, were amplified via PCR using specific I-IV primers (**Supplementary Table 1**) and plasmid pRS300 as template. The artificial miRNA precursors were digested with *Spe*I and *Kpn*I and inserted into the pMDC32 binary vector. All the constructs were identified via DNA sequencing.

Reverse Transcription-Polymerase Chain Reaction (RT-PCR) and Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

For RT-PCR, total RNA was isolated using the RNAiso Plus reagent (TaKaRa)[2]. First-strand cDNA was synthesized using 500 ng of total RNA and the one-step RT-PCR kit (TaKaRa). PCR fragments were subsequently amplified using their corresponding primers (**Supplementary Table 1**), analyzed via agarose gel electrophoresis, and stained with the Goldview^{TM} nucleic acid stain (SBS Genetech Co., Ltd., China).

We then conducted qRT-PCR. Total RNA (500 ng) isolated from the leaves was reverse transcribed using the SYBR PrimeScript^{TM} RT-PCR Kit (TaKaRa) to synthesize cDNA. PCR amplification was performed using the SYBR® Premix Ex Taq^{TM} kit (TaKaRa). The gene-specific primers used are listed in **Supplementary Table 1** for the qRT-PCR reactions. *eIF4A1* was used as an internal control.

Total Protein Extracts and Western Blot Assay

Total proteins were extracted from 1 g of 14-day-old seedlings and dissolved in sample buffer (50 mM Na_2HPO_4/NaH_2PO_4, pH 7.4; 150 mM NaCl; 1% Triton X-100; 15% glycerol; 1 mM PMSF; and 1 × cocktail). Isolated proteins were identified using 10% sodium dodecyl (lauryl) sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. They were transferred onto polyvinylidene difluoride (PVDF) membranes using a semi-dry electroblotter (Bio-Rad). The PVDF membranes were probed with anti-FLAG (Sigma), anti-H3 (Agrisera), anti-LFR (Lin et al., 2018), anti-SWI3B (Sarnowski et al., 2002) or anti-tubulin antibody (Sigma). Goat anti-rabbit or anti-mouse IgG secondary antibodies were used for immunodetection.

Co-immunoprecipitation (co-IP) Assay

Approximately 4 g of 10-day-old *Arabidopsis* seedlings were used for immunoprecipitation experiments. The seedlings were extracted and added to a 4 mL protein solution buffer (50 mM Na_2HPO_4/NaH_2PO_4, pH 7.4; 150 mM NaCl; 1% Triton X-100; 15% glycerol; 1 mM PMSF; and protease inhibitor cocktail from Roche). The extracts were centrifuged at 17,000 × *g* for 10 min at 4°C. The supernatant proteins were then incubated with 40 μL of anti-FLAG M2 agarose beads (Sigma, Cat. # M8823) for 1 h at 4°C. After incubation, the beads were collected by centrifugation and washed three to five times with 1 mL wash buffer (50 mM Na_2HPO_4/NaH_2PO4, pH 7.4; 150 mM NaCl; 0.1% Triton X-100; 10% glycerol; 1 mM PMSF; and protease inhibitor cocktail from Roche). The antigen-antibody complex was boiled in Laemmli SDS-PAGE buffer (125 mM Tris–HCl, pH 6.8; 4% SDS; 20% glycerol; 2% mercaptoethanol; and 0.001% bromophenol blue), separated on a 12% SDS-PAGE gel, and transferred onto a PVDF membrane. Proteins immunoprecipitated with the anti-FLAG antibodies were probed with anti-LFR polyclonal antibody, which was previously prepared in our lab (Gao et al., 2008), or with anti-SWI3B antibody reported previously (Sarnowski et al., 2002). Secondary antibody (goat anti-rabbit IgG) was used for immunodetection. The SuperSignal West Femto System (Pierce) was used for signal detection.

Yeast Two-Hybrid (Y2H) Analysis

The Y2H screening of cDNA library derived from 9-day-old seedlings of Arabidopsis was performed following the manufacturer's instructions (Matchmaker GAL4 Two-Hybrid System 3 & Libraries User Manual Clontech Laboratories). We conducted a Y2H assay. Briefly, the bait plasmid, pGBKT7, or prey plasmid, pGADT7, with full-length or truncated LFR or SWI3B were co-transformed into AH109. The co-transformed colonies were selected to grow on a selective medium that lacked leucine and tryptophan (SD/-L-W). A growth assay was then conducted, in which the physical interaction between different pair of proteins was tested on selective medium that lacked leucine, tryptophan, adenine, and histidine (SD/-L-W-A-H). Liquid β-galactosidase (β-Gal) assays, with o-nitrophenyl β-D-galactopyranoside (ONPG) (Sigma) as a substrate, were measured as described in the manufacture's handbook (Clontech Yeast Protocols Handbook). One unit of β-galactosidase activity was defined as the amount in which hydrolysis of 1 μ mol of ONPG to o-nitrophenol and D-galactose per min per cell occurred.

Bimolecular Fluorescence Complementation (BiFC) Assay

The BiFC assay was performed as previously described (Ou et al., 2011). The plasmids were separately introduced into Agrobacterium GV3101 and co-infiltrated into the young flattened leaf blade of *Nicotiana benthamiana*. After incubation for approximately 48 h, images were captured using a Zeiss LSM

[1]http://wmd3.weigelworld.org/cgi-bin/webapp.cgi

[2]http://www.clontech.com/takara

710 confocal microscope. Green fluorescent protein (GFP) and 4, 6-diamidino-2-pheylindole (DAPI) signals were examined at 488 and 405 nm, respectively.

Chromatin Immunoprecipitation (ChIP) Assay

The chromatin immunoprecipitation and qPCR (ChIP-qPCR) assay was carried out as previously described (Yamaguchi et al., 2014) with minor modifications. Approximately 0.3–0.6 g of seedlings or the first to third rosette leaves of the 14-day-old seedlings were crosslinked with 1% formaldehyde and fully ground in liquid nitrogen. Chromatin was isolated and cut into approximately 500 bp DNA fragments via sonication. The chromatin suspension was incubated for 2 h with 50 μL of magnetic protein G beads (Invitrogen, Cat. # 10004D), 5 μg of anti-FLAG antibody (Sigma Cat. # F3165), 5 μg of anti-trimethyl-histone H3 (Lys4) (Millipore Cat. # 07-473), or 2 μL of anti-LFR rabbit polyclonal antiserum from our laboratory (Gao et al., 2008). Pre-immune serum was used as the control. DNA was isolated using the DNA purification kit (Qiagen, Cat. # 28104) and used as the template of primers listed in **Supplementary Table 1** in real-time qPCR.

RESULTS

LFR Physically Interacts With SWI3B in Yeast and Plant

In our previous study, the transcription factor *AS2* was identified as an *LFR*-interacting protein through genetic screening, which explained the molecular mechanism of the developmental of defects in petiole length, vasculature pattern, and leaf margin except for the leaf blade upward-curling defects in *lfr* mutants (Lin et al., 2018). To further elucidate the molecular mechanisms of the functions of LFR in flattened leaf blades, we screened the cDNA library of 9-day-old seedlings of Arabidopsis to identify possible LFR-interacting proteins by Y2H. Since the full-length LFR had transcriptional self-activation activity in the yeast AH109 strain (Yuan et al., 2012), LFRΔC2 (1–310 amino acids), which has no transcriptional self-activation activity, was used as a bait to screen the cDNA library. A total of 79 positive in-frame proteins were identified (**Supplementary Table 2**). Sequencing analysis showed that one positive colony contained the full-length coding sequence of *SWI3B*, which was a component of the SWI/SNF complex in Arabidopsis. To further verify the interaction between LFR and SWI3B, the Y2H assay was performed using the full-length LFR, which was fused with AD and BD-SWI3B. Yeast AH109 colonies, which were co-transformed with BD-SWI3B and AD-LFR, grew well on selective medium and had a much higher β-Gal activity than the negative control; however, BD-SWI3B and AD-LFR had no self-activation (**Figure 1A**). These findings indicate that LFR interacts directly with SWI3B in yeast.

To further confirm the interaction between LFR and SWI3B in plant cells, BiFC assay was performed in *N. benthamiana* leaves. We observed GFP signals in cells that were co-transformed with *cCFP-LFR/nYFP-AS2* plasmid as a positive control (Lin et al., 2018), but GFP signals were rarely observed in nuclei that were co-transformed with cCFP empty vectors, nYFP-SWI3B or nYFP empty vectors, and cCFP-LFR (**Figure 1B**). Under these experimental conditions, GFP signals were observed in the nuclei of epidermal cells co-transformed with cCFP-LFR and nYFP-SWI3B (**Figure 1B**). Therefore, the results of the BiFC assay show that LFR interacts with SWI3B in plant.

To further test whether LFR interacts with SWI3B in Arabidopsis, we prepared transgenic complementary lines of *35S:LFR-3FLAG/lfr-1* and *35S:SWI3B-3FLAG/swi3b-2* (**Supplementary Figures 1**, **2**). We performed a co-immunoprecipitation (co-IP) assay using total protein extracts isolated from LFR-3FLAG or SWI3B-3FLAG transgenic seedlings. Anti-FLAG antibody-coated beads were used to immunoprecipitate LFR-3FLAG and its associated proteins. We then used anti-SWI3B antibodies to detect endogenous SWI3B proteins, which were only detected in *35S:LFR-3FLAG/lfr-1* transgenic rescue plants but not in their wild-type counterparts (**Figure 1C**). In the co-IP assay in *35S:SWI3B-3FLAG* transgenic rescue plants, SWI3B-3FLAG was also specifically co-precipitated with endogenous LFR (**Figure 1D**). These co-IP data indicate that LFR and SWI3B co-exist in the same complex in Arabidopsis.

There are four SWI3 proteins in the genome of Arabidopsis: SWI3A, SWI3B, SWI3C, and SWI3D (Sarnowski et al., 2005). We examined the interactions between the following pairs in yeast: LFR and SWI3A; LFR and SWI3C; and LFR and SWI3D. Since BD-SWI3C and BD-SWI3D had transcriptional self-activation activity, yeast AH109 colonies co-transformed with BD-LFRΔC2 and AD-SWI3C/SWI3D were tested via a growth assay. The Y2H results show that in yeast, LFR interacts with SWI3A but not with SWI3C and SWI3D (**Supplementary Figure 3**).

The N-Terminal Domain of LFR Is Essential for the Interaction Between LFR and SWI3B

Our previous report indicated that the C terminus of LFR had three predicted protein–protein interaction ARM-repeat domains responsible for the self-activation activity of BD-LFR in yeast (Yuan et al., 2012). Here, a series of truncated LFR without self-activation activity was used to further identify the interaction domain between LFR and SWI3B in yeast. The truncated LFR without the ARM domains, including LFRΔC1-C4, interacted with SWI3B (**Figures 2A,B**, **Supplementary Figure 4**). Upon deletion of the N-terminal domain (ND) of LFR (25–60 amino acids), the interaction between LFR and SWI3B was abolished (**Figures 2A,B** and **Supplementary Figure 4**), indicating that the ND motif of LFR was essential for its interaction with SWI3B.

We also determined the specific region of SWI3B that was involved in the interaction with LFR. The SWI3B protein included the SWIRM domain, zinc finger (ZF, homologous with the ZF domain of SWI3D), SANT, and leucine zipper (LZ) domain (Bateman et al., 1999; Sarnowski et al., 2005). The results

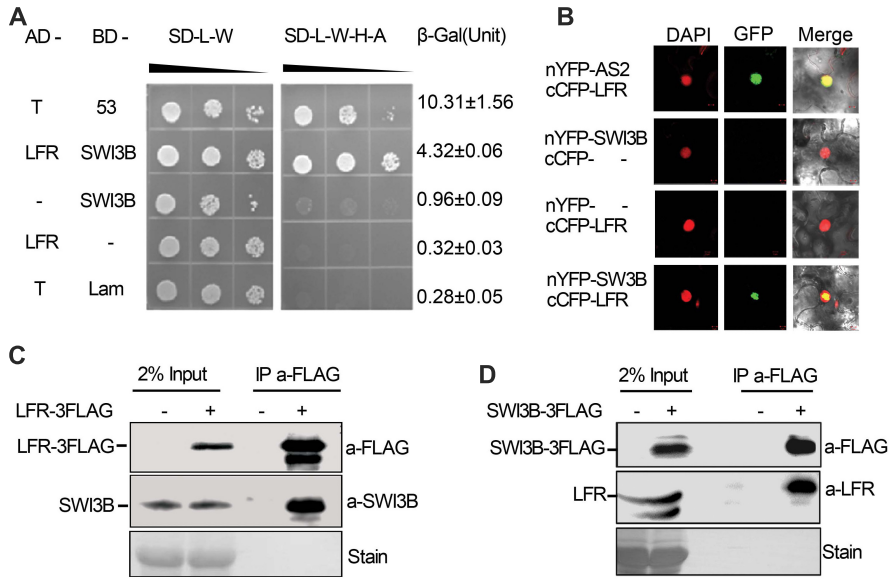

FIGURE 1 | Leaf and flower related (LFR) interacts with SWI3B in yeast and *planta*. **(A)** The growth assay and quantitative β-galactosidase (β-Gal) activity assays showing that AD-LFR interacts with BD-SWI3B in the Y2H assay. The growth experiment was performed on selective medium (SD/-L-W and SD/-L-W-A-H) after gradient dilution (10^{-1}, 10^{-2}, and 10^{-3}) as indicated by black triangles. The AD-T/BD-p53 and AD-T/BD-Lam co-transformed yeast colonies were used as the positive and negative control, respectively. In quantitative β-Gal activity assays, data are mean ± standard error from three independent experiments. **(B)** BiFC assay showing that cCFP-LFR interacts with nYFP-SWI3B in transiently transformed epidermal cells of tobacco leaf (22 of 114 cells had GFP signal). DAPI signal indicates nucleus. GFP signal shows interaction. Merge means overlay of DAPI and GFP fluorescence signals. cCFP-LFR/nYFP-AS2 serves as a positive interaction control; - no protein fusion. **(C,D)** Co-IP assay identifies LFR-3FAG and SWI3B co-exist in *35S:LFR-3FLAG/lfr-1* transgenic rescue line **(C)**, and SWI3B-3FAG and LFR co-exist the same complex in *35S:SWI3B-3FLAG/swi3b-2* transgenic rescue line **(D)**. Total protein extracts were derived from 14-day-old seedlings of Col-0 (–) or transgenic rescue line (+). Anti-FLAG antibody beads were used to immunoprecipitate (IP a-FLAG). In western blot, anti-FLAG (a-FLAG) or anti-SWI3B (a-SWI3B) or anti-LFR (a-LFR) antibody was used to detect LFR-3FLAG/SWI3B-3FLAG or native SWI3B or LFR, respectively. Ponceau stain (stain) serves as the loading control.

of the growth assay reveal that all combinations, except for BD-ZF and AD-LFR, can activate the reporter genes (**Figures 2C,D** and **Supplementary Figure 4B**). These results suggest that SWIRM, SANT, and LZ but not the ZF domain of SWI3B were able to interact with LFR.

The Biological Function Analysis of Truncated LFR by Transgenic Rescue Assay

To explore the importance of the ND motif for the biological function of LFR in plant development, we fused *LFRΔN1* and *LFRΔN2* with *3FLAG* driven by the *CaMV 35S* promoter to obtain *35S:LFRΔN1-3FLAG* and *35S:LFRΔN2-3FLAG*, respectively. We then transformed them into the *lfr-2* background (**Figure 3A**). As a control, *35S:LFR-3FLAG* completely rescued the upward-curling leaf and sterile defects of *lfr-2*. Four transgenic lines of *35S:LFRΔN2-3FLAG/lfr-2* without the N-terminal 1–25 amino acids could also recover the defects of *lfr-2* in leaf and silique development. However, the *35S:LFRΔN1-3FLAG* construct, with further deletion of the ND region of LFR, could not rescue any phenotype of *lfr-2* (**Figure 3B**). To ensure that *LFRΔN1-3FLAG* was normally expressed, we carried out RT-PCR and Western blotting and found that it could be expressed normally at both the RNA and protein levels (**Figures 3C,D**). These data demonstrate that the ND motif responsible for the LFR-SWI3B interaction is crucial for the full biological function of LFR in Arabidopsis.

Meanwhile, we constructed a truncated LFR without one or more ARM domains, including *LFRΔC1*, *LFRΔC3*, and *LFRΔC4*, into the *35S:3FLAG* vector (**Figure 3A**). We then introduced them into the *lfr-2* background. Three transgenic lines of *35S:LFRΔC3-3FLAG/lfr-2* and six transgenic lines of *35S:LFRΔC4-3FLAG/lfr-2* were observed to partially rescue the leaf and silique phenotypes of *lfr-2* (**Figure 3B**). However, the transgenic homozygous lines expressing *35S:LFRΔC1-3FLAG* without the ARM 1–3 domain could not rescue any phenotype of *lfr-2* (**Figure 3B**). In addition, we conducted RT-PCR and Western blotting and found that *LFRΔC1-3FLAG* was normally expressed at both the RNA and protein levels (**Figures 3C,D**).

Together, these transgenic complementary data suggest that both the ND and ARM domains are crucial for the biological function of LFR in Arabidopsis.

LFR Genetically Interacts With *SWI3B* During Leaf Blade Development

To detect the genetic relationship between *LFR* and *SWI3B*, we created the knock-down mutants of *SWI3B* using artificial microRNA to produce *SWI3B-amic* because the null mutants of *SWI3B* (*swi3b-1* and *swi3b-2*) were embryo-lethal (Sarnowski et al., 2005). We chose two sites, *mic1* and *mic2* (short for

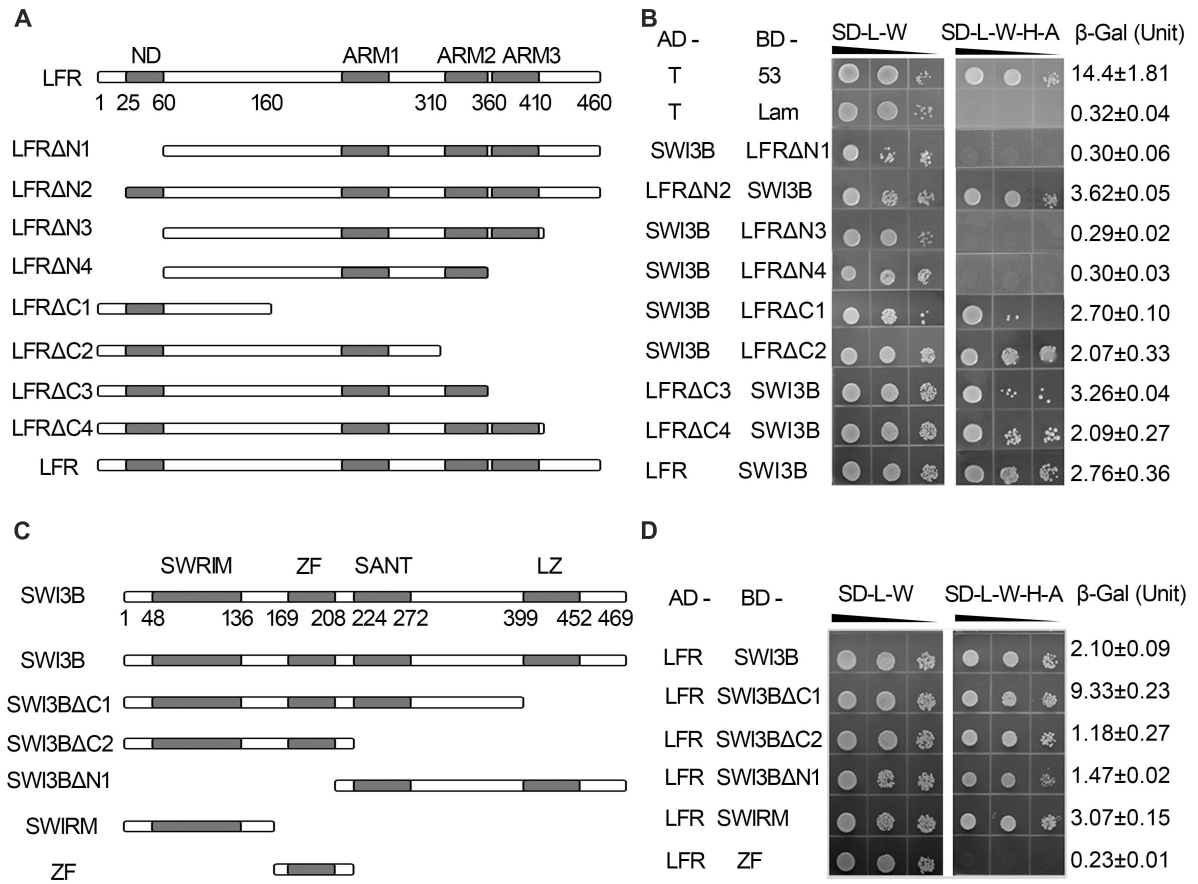

FIGURE 2 | The interacting domain analysis of LFR and SWI3B. **(A,C)** Schematic of full-length and truncated LFR **(A)** and SWI3B **(C)**. **(B,D)** The interacting domain of LFR with full-length SWI3B **(B)** and interacting domain of SWI3B with full-length LFR **(D)** used in Y2H growth and quantitative β-Gal activity assay. Yeast colonies were tested for growth assay on SD-L-W or SD-L-W-H-A after gradient dilution (10^{-1}, 10^{-2}, and 10^{-3}) as indicated by black triangles. Numbers on the right represent the mean ± standard error of three biological replicates of the β-Gal activity.

SWI3B-amic1 and *SWI3B-amic2*), for the design of *SWI3B* artificial microRNA (**Figure 4A**). We obtained four independent transgenic homozygous lines for *mic1* and ten for *mic2* in the wild-type background. Transgenic *mic1-2* and *mic2-6* lines, which had low transcript level of *SWI3B*, had upward-curling leaves compared to the wild type (**Figures 4B,C**). To detect whether *mic1-2* and *mic2-6* specifically targeted *SWI3B*, we measured the expression levels of *SWI3B* homologous genes, which included *SWI3A*, *SWI3C*, and *SWI3D*. The transcript levels of *SWI3C* and *SWI3D* had no obvious changes; however, a slight increase in the *SWI3A* transcript was noted (**Figure 4D**). These data indicate that *mic1-2* and *mic2-6* specifically target *SWI3B* and result in an upward-curling leaf phenotype, which is similar to that of plants with *LFR* loss-of-function mutation (Wang et al., 2009; **Figure 3B**).

We then obtained double mutants of *lfr-1/2 mic1-2* or *lfr-1/2 mic2-6* by genetic crossing. qRT-PCR data showed that the double mutants had significantly reduced the expression of *LFR* and *SWI3B* (**Figure 5A**). The transcript and protein levels of *SWI3B* did not change significantly in *lfr* mutants (**Figures 5A,B**). Meanwhile, we did not detect obvious changes in LFR at the RNA and protein levels in *mic1-2* and *mic2-6* (**Figures 5A,C**). Therefore, these data indicate that LFR and SWI3B do not regulate each other at the transcriptional and protein levels. We then analyzed the phenotypic characteristics of *lfr-1*, *lfr-2*, *mic1-2*, and *mic2-6* single and double mutants. *lfr-1*, *lfr-2*, *mic1-2*, and *mic2-6* all displayed upward-curling leaves and had a sawtooth appearance at the margin of the leaf blade (**Figures 5D,E**). The upward-curling leaf phenotype in double mutant *lfr-1 mic1-2* or *lfr-1 mic2-6* was a little stronger that of the single mutant (**Figure 5D**). The same results were observed in the *lfr-2 mic1-2* or *lfr-2 mic2-6* double mutants (**Figure 5E**).

Taken together, these genetic data suggest that *LFR* and *SWI3B* may have overlapping functions in the regulation of flattened leaf blade development.

The Differentially Expressed Genes in *lfr-2* and *SWI3B-amic* Leaves

To identify differentially expressed genes in *lfr-2* and *SWI3B-amic* leaves, we examined the transcript levels of genes encoding the major transcription factors involved in the control leaf polarity,

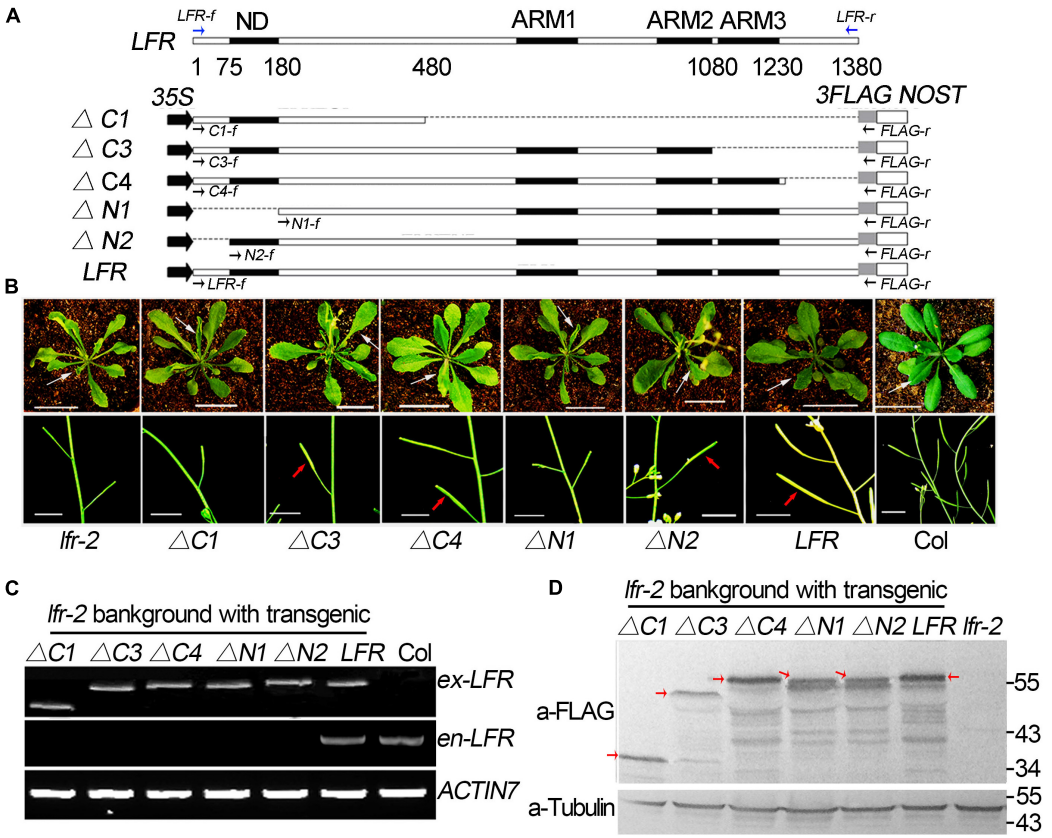

FIGURE 3 | Transgenic rescue analysis of truncated LFR in *lfr-2* null mutant. **(A)** Schematic map of constructs used in the transgenic rescue analysis. *LFR* indicated full length coding sequence (CDS) of LFR (1380 bp). *35S*, *NOST*, and *3FLAG* represented the *CaMV 35S* promoter, the terminator sequence of the *NOS* gene, and the tag, respectively. The numbers below LFR gene structure showed the exact nucleic acids of the LFR CDS coding for the corresponding protein domains. Full length and truncated LFR (*LFR*, Δ*C1*, Δ*C3*, Δ*C4*, Δ*N1*, and Δ*N2*) were inserted into the *35S:3FLAG* vector and transformed into *lfr-2*. The blue and black arrows represent primers used for endogenous *LFR* (*en-LFR*) and different lengths of exogenous *LFR* (*ex-LFR*), respectively, used in RT-PCR in panel **(C)**. **(B)** The leaf (from 35-day-old plants, upper panel) and siliques (from 50-day-old plants, bottom panel) of Col, *lfr-2*, and different transgenic lines in the *lfr-2* background. The white arrowhead pointed to the leaves from a similar position of different genotypes. The red arrowhead pointed to the elongated siliques with seeds. Bar = 2 cm in upper panel, Bar = 1 cm in bottom panel. **(C)** RT-PCR analysis of endogenous (*en-*) and exogenous (*ex-*) full length or truncated *LFR* in different genotypes with the primers showed in panel **(A)**. *ACTIN7* was used as the loading control. **(D)** Western blot with anti-FLAG monoclonal antibody (a-FLAG) or anti-Tubulin (a-Tubulin) in Col and transgenic lines. The red arrows represented the corresponding truncated or full LFR-FLAG fusion proteins. The signal underlying LFR-FLAG fusion protein is caused by the degradation of LFR-FLAG fusion proteins. Tubulin was used as the internal loading control.

including *HD-ZIP III* (*PHABULOSA*, *PHB*; *PHAVOLULA*, *PHV*; *REVOLUTA*, *REV*) for adaxial cell fate determination, and *YAB1* (*FIL*) and *KAN* (*KAN1* and *KAN2*) family genes for abaxial cell fate establishment. And the *ASYMMETRIC LEAVES2* (*AS2*) and Knotted in *A. thaliana* (*KNAT*) and some other genes which were already tested previously in our study (Lin et al., 2018) were not included here. We also examined several genes related to auxin metabolism and synthesis, including *IAMT1*, *INDOLE-3-ACETIC ACID INDUCIBLE 17* (*IAA17*), *IAA3*, and *YUCCA* (*YUC6*) in the wild type, *lfr-2*, and *SWI3B-amic* single and double mutants. There was a significant increase in the expression levels of *IAMT1* in *lfr-2* and *SWI3B-amic* single mutants and even higher transcription levels in the double mutants (**Figure 6A**). In addition, *YUC6* was also significantly increased at the transcriptional level in the leaves of the *SWI3B-amic* mutants compared to that in the wild type. However, there was no significant change in *YUC6* in the *lfr-2* mutant. Furthermore, the double mutants had a similar expression to that of the *lfr-2* mutant (**Figure 6A**). In addition, the abaxial gene, *FIL*, was decreased at the transcriptional level in the leaves of *lfr-2* and *SWI3B-amic* single and double mutants compared to that in the wild type (**Figure 6B**). However, there were no significant changes in the *HD-ZIP III* and *KAN* family genes. These results show that LFR and SWI3B play similar roles in the transcriptional regulation of the expression of *IAMT* and *FIL* in Arabidopsis leaves.

LFR and SWI3B Are Enriched in Chromatins of *FIL* and *IAMT1*

Since we found that the expression of *FIL* was downregulated in both *SWI3B-amic* and *lfr-2* (**Figure 6B**), we speculated that SWI3B might be a partner of LFR in regulating *FIL* expression. First, we tested whether LFR was tethered to the

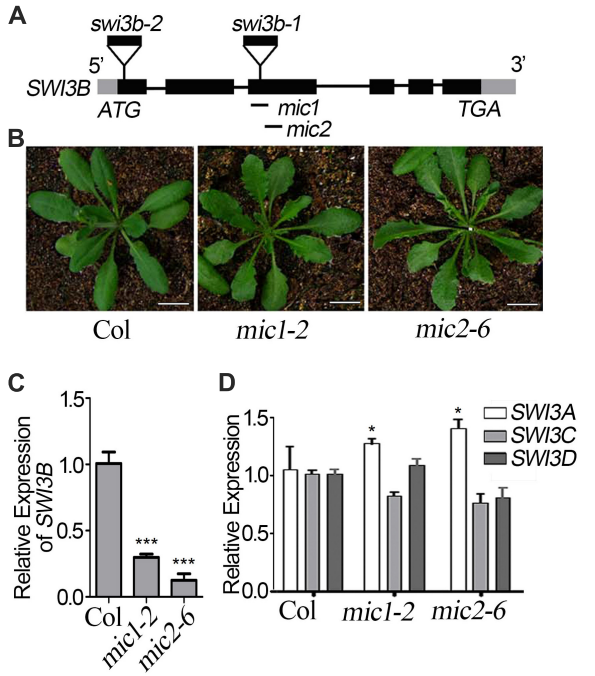

FIGURE 4 | Transgenic lines of artificial miRNA-mediated knockdown of *SWI3B*. **(A)** Schematic gene model and mutants of *SWI3B*. *swi3b-1* and *swi3b-2* are T-DNA insertion mutants reported previously (Sarnowski et al., 2005). The target sites of two artificial microRNA of *SWI3B* (*mic1*: 594–614 bp and *mic2*: 753–773 bp) were indicated by black lines below the gene structure. **(B)** The phenotypes of T1 transgenic lines of artificial microRNA of *SWI3B* (*mic1* and *mic2*). The appearance of 35-day-old plants from different genotypes as indicated. Scale bars = 1 cm. **(C,D)** SWI3B genes expression level **(C)** and *SWI3A*, *SWI3C*, and *SWI3D* genes expression level **(D)** analyzed by quantitative reverse transcription-PCR (qRT-PCR). The RNA was extracted from the 35-day-old Col and *SWI3B-amic* T1 lines. *eIF4A1* was used as an internal control. Bars indicate the means ± SE of three independent biological repeats. Significant statistical differences were tested using Student's *t*-test (***$P < 0.001$, *$P < 0.05$).

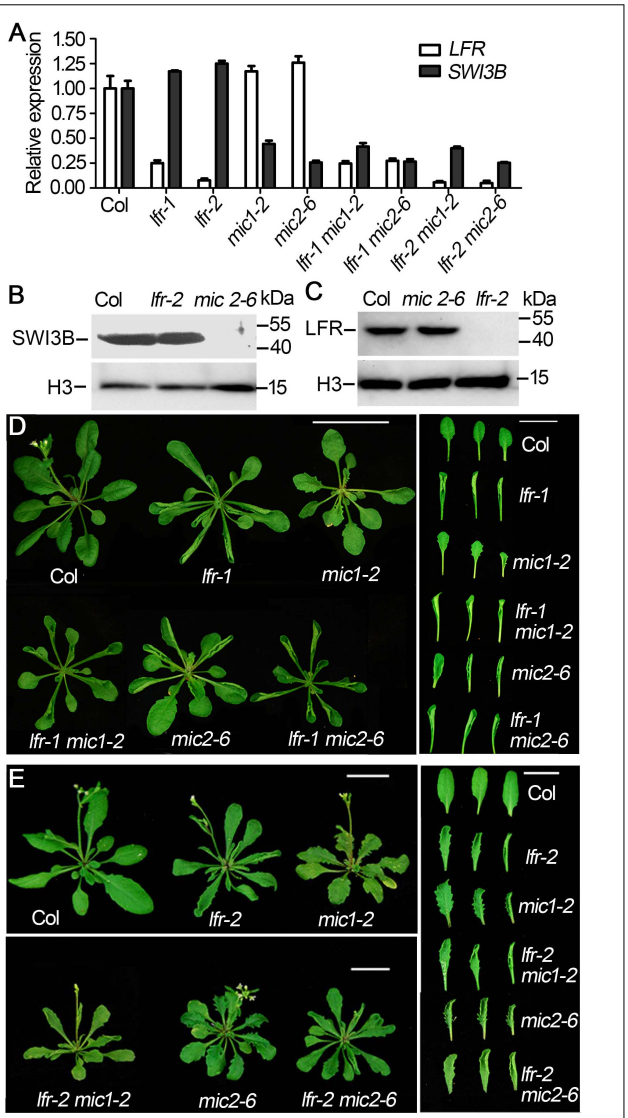

FIGURE 5 | The genetic interaction between *lfr* and *SWI3B-amic* in leaf development. **(A)** qRT-PCR analysis of *SWI3B* and *LFR* transcript in 35-day-old Col and *lfr-1*, *mic1-2*, and *mic2-6* single or double mutant plants. The total RNA was isolated from the 7th and 8th rosette leaves of 35-day-old Col or various mutants. *eIF4A1* was used as an internal control. **(B,C)** Western blot analysis of SWI3B **(B)** and LFR **(C)** protein level in from different genotypes as indicated, and the H3 served as a loading control. The total protein was extracted from the 14-day-old seedlings of Col and different mutants. **(D,E)** Genetic interaction between *SWI3B-amic* and *lfr-1* were assessed at 35 days of age **(D)** and *lfr-2* at 40 days of age **(E)**. The phenotype of whole plants (left) and the representative 8th–10th rosette leaves (right) from different genotypes as indicated. Scale bars = 2 cm in panels **(D,E)**.

FIL locus by conducting a ChIP-qPCR assay. The upstream *b-c* fragments of the *FIL* promoter were reproducibly amplified from the chromatin of *LFR:LFR-FLAG/lfr-2* immunoprecipitated with anti-FLAG; however, no enrichment was detected in Col (**Figures 7A,B**). To determine whether SWI3B was also tethered to the *FIL* locus, we performed ChIP-qPCR in *35S:SWI3B-FLAG/swi3b-2* transgenic plants, and the significant enrichment of SWI3B-FLAG at *b-c* fragments of *FIL* chromatin was reproducibly detected in SWI3B-FLAG fusion protein compared to that in the Col control (**Figure 7C**). These results suggest that there is an association between LFR and SWI3B and the *FIL* promoter, thereby indicating that *FIL* is the direct target gene of LFR and SWI3B.

To further investigate whether the binding activity of LFR to the *FIL* locus was dependent on SWI3B, we performed a ChIP-qPCR assay in *mic2-6* mutant plants using anti-LFR antibodies. In the absence of functional *SWI3B*, the enrichment of LFR at fragments *b* and *c* of the *FIL* promoter was partly reduced compared to that in the wild type (**Figure 7D**). To rule out the possibility that the reduction in binding ability might result from low LFR levels in *mic2-6*, we conducted Western blotting and found that the protein level of LFR in *mic2-6* was almost comparable to that in the wild-type control (**Figure 5C**). These data indicated that LFR and SWI3B co-target the *FIL* locus. Furthermore, the binding of LFR to the *FIL* locus is partly dependent on SWI3B.

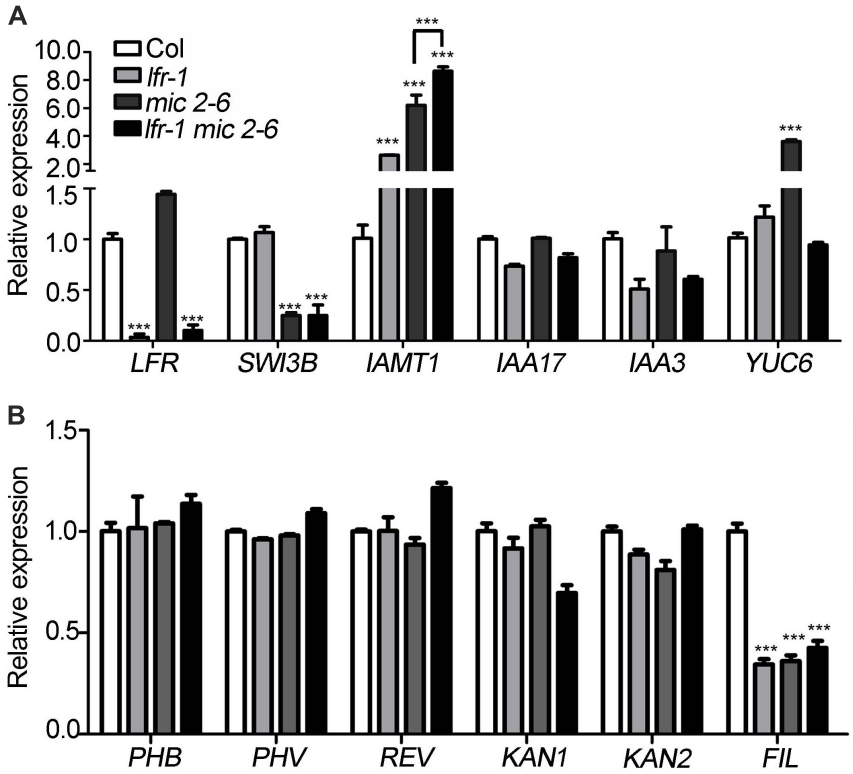

FIGURE 6 | Differentially expressed genes in *lfr-1*, *mic2-6* single and double mutants. **(A,B)** The qRT-PCR data for tested the transcript level of auxin metabolism and synthesis genes **(A)** and leaf polarity genes **(B)** in different backgrounds as indicated. The total RNA was isolated from the 7th and 8th rosette leaves of 35-day-old Col or various mutants. Transcript levels were normalized to loading control gene *eIF4A1*. Bars indicate the means ± SE of three independent biological repeats. Significant statistical differences were tested using Student's *t*-test (***$P < 0.001$).

In addition, it was reported that the overexpression of *IAMT1* caused upward-curling leaf in *SWI3B-RNAi* plants (Han et al., 2018), but it is not clear that whether SWI3B was associated with the *IAMT1* chromatin. Since *IAMT1* transcript levels were increased in *lfr-2* and *SWI3B-amic* plants (**Figure 6A**), we tested the association between SWI3B and LFR and the chromatin of *IAMT1*. ChIP-qPCR assay data showed that fragment 4 (−65 to 45) was reproducibly amplified from the chromatin of *LFR:LFR-FLAG/lfr-2* or *35S:SWI3B-FLAG/swi3b-2* transgenic plants immunoprecipitated with anti-FLAG. However, no enrichment was detected in Col (**Figures 7E–G**). However, we did not detect any enrichment signals of LFR or SWI3B at the *YUC6* locus (**Supplementary Figure 5**). Moreover, the enrichment of LFR in the chromatin of *IAMT1* was partly reduced in the *mic2-6* mutant compared to that in the wild type (**Figure 7D**). These results indicate that LFR and SWI3B co-target the *IAMT1* locus *in vivo*. Moreover, the binding activity of LFR to the *IAMT1* locus is partly dependent on SWI3B.

Increased *FIL* Expression Partially Recovers Upward-Curling Leaf Phenotype of *lfr* Mutant

To further establish the link between *FIL* expression and the upward-curling leaf phenotype of *lfr*, we conducted the genetic analysis by introducing *35S:FIL* into *lfr-2* heterozygous plant background. The *FIL* expression levels were increased by different degrees in the transgenic lines, *35S:FIL 2-1-7* and *35S:FIL 17-8-20* in both the wild type (WT) and *lfr-2* background (**Figure 8A**). Though the rosette leaves of *35S:FIL 2-1-7*/WT displayed largely the same morphology as those of the Col, the *35S:FIL 17-8-20*/WT exhibited obviously downward-curling leaf phenotype, which may be resulted from the significant overexpression of *FIL* (**Figures 8A,B**; Bonaccorso et al., 2012). Intriguingly, the increased expression level of *FIL* can partially recover the upward-curling leaf phenotype of *lfr-2* (**Figure 8B**). These results indicated that the downregulation of *FIL* may be one of the possible causes of the upward-curling leaf phenotype of *lfr*, which provides genetic evidence for the regulation of *FIL* by LFR. Besides, we also found that the double mutants of *35S:FIL 2-1-7*/*lfr-2* and *35S:FIL 17-8-20*/*lfr-2* had smaller and more leaves than the Col, indicating that there might be some phenotype enhancement in the process of SAM development when overexpressing *FIL* in the *lfr*-2 background.

DISCUSSION

Our previous study demonstrated that Arabidopsis LFR plays pivotal roles during leaf and flower development (Wang et al.,

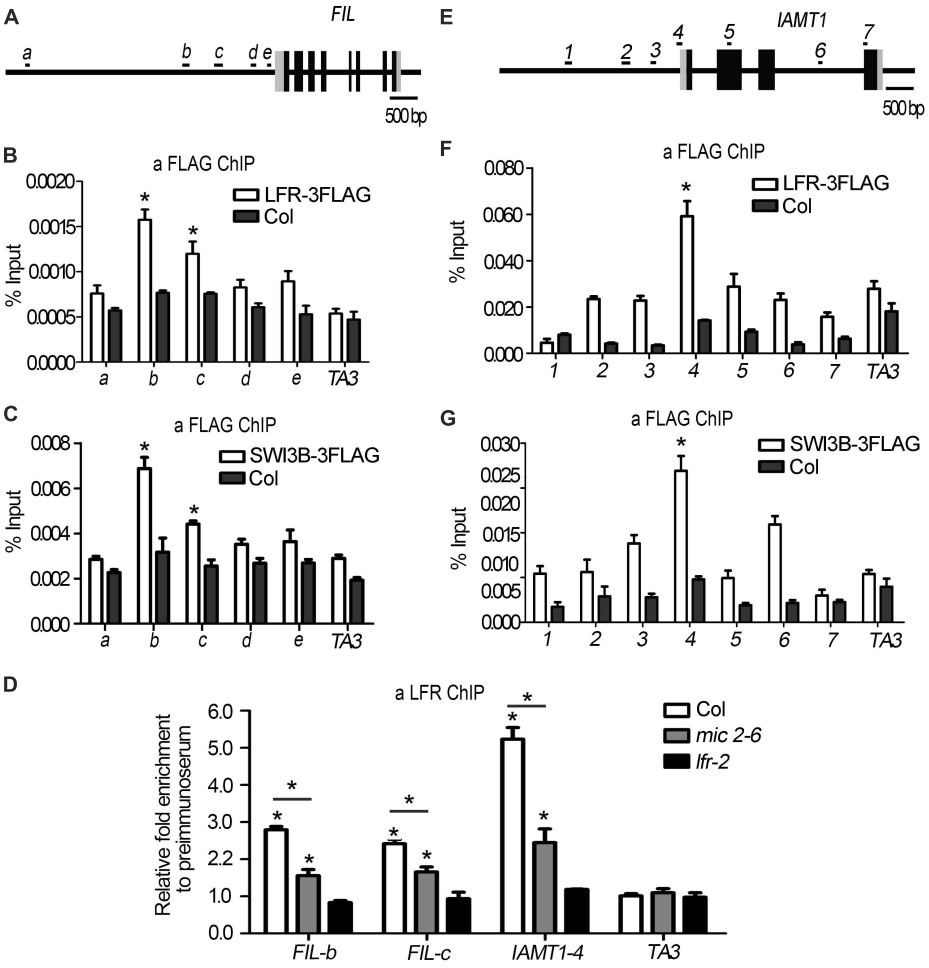

FIGURE 7 | LFR and SWI3B are associated with the chromatin of *FIL* and *IAMT1*. **(A,E)** The diagrams of *FIL* and *IAMT1* gene structures. The black boxes indicate exons, the gray boxes indicate untranslated regions and the long black lines represent the upstream sequence or promoter, introns regions, or 3′-terminal sequence. The lowercase letters **(A)** or the numbers **(E)** and black short lines above the gene structures represent PCR fragments tested in ChIP-qPCR **(B–D,F,G)**. **(B,F)** ChIP-qPCR assay to test the association of LFR-3FLAG with *FIL* **(B)** and IAMT1 **(F)** chromatin using anti-FLAG antibody. **(C,G)** ChIP-qPCR assay to test the association of SWI3B-3FLAG with *FIL* **(C)** and IAMT1 **(G)** locus using anti-FLAG antibody. **(D)** ChIP-qPCR assay to test the association of LFR to *FIL* chromatin using the anti-LFR antibody in *mic2-6*. The bars represent the means of three independent biological repeats and the error bars stand for SE. Significant statistical differences were tested by Student's *t*-test (*$P < 0.05$). A retrotransposon locus *TA3* (*At1g37110*) was used as the negative control in ChIP-qPCR **(B–D,F,G)**.

2009, 2012; Lin et al., 2018). LFR encodes a nuclear protein with ARM-repeat domains (Wang et al., 2009). Through genetic screening, we identified that LFR synergistically interacts with AS2 to repress *BP* expression in the specific processes of leaf development, such as leaf petiole length, the formation of leaf midvein, and elongation of leaflet-like structure at the leaf margin (Lin et al., 2018). However, LFR and AS2 seem to act oppositely in control of the flattened leaf development. To further elucidate the molecular mechanism of LFR in flattened leaf blade development, we isolated the SWI/SNF complex subunit, SWI3B, as another interacting partner of LFR by Y2H screening. This interaction was confirmed by BiFC and co-IP (**Figure 1**). Y2H and transgenic complementary assays of different truncated LFR proteins showed that the ND domain of LFR was essential for its interaction with SWI3B and was important for its biological function in Arabidopsis (**Figures 2**, **3**). Consistent with a previous report (Han et al., 2018), the knock-down mutant of *SWI3B* by artificial miRNA (*SWI3B-amic*) resulted in an upward-curling leaf phenotype, which was similar to those of the *lfr-1* and *lfr-2* mutations (**Figure 4**). Different combinations of *lfr-1/lfr-2* and *mic1-2/mic2-6* double mutants also exhibited upward-curling leaves to a little stronger than the single mutants (**Figure 5**). In addition, LFR and SWI3B co-targeted similar chromatin regions of *FIL* and *IAMT1*, which were differentially expressed in the single and double mutants of *lfr-2*, *SWI3B-amic*, and double mutants (**Figures 6**, **7**). Furthermore, the association between LFR and *FIL* or *IAMT1* was partly dependent on SWI3B (**Figure 7D**). Interestingly, we notice that the expression of *IAMT1* in *lfr SWI3B-amic* double mutant is higher than either single mutant (**Figure 6A**), and the curly leaf phenotype of *lfr SWI3B-amic* double mutant seems a little stronger than single mutants (**Figures 5D,E**), suggesting that besides the

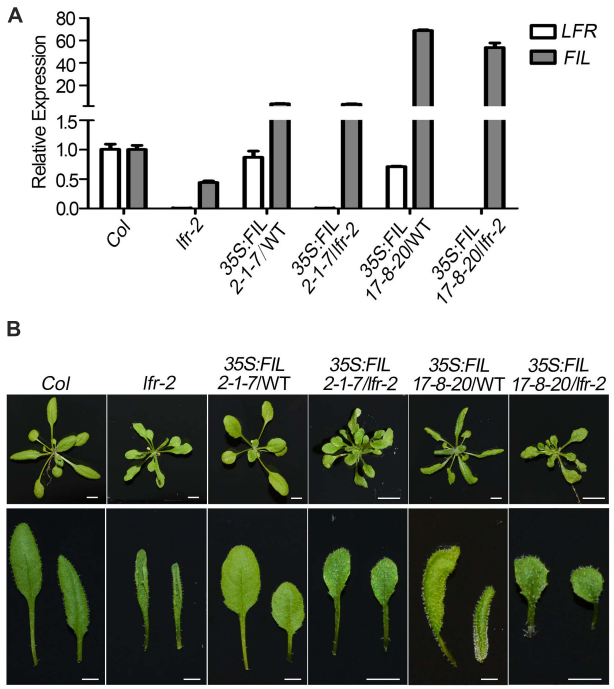

FIGURE 8 | Overexpression of *FIL* partially rescued the upward-curling leaf defect of *lfr*. **(A)** The qRT-PCR data to detect the transcript level of *FIL* in different backgrounds as indicated. The total RNA was isolated from the rosette leaves of 22-day-old Col, *lfr-2*, or the transgenic plants. Transcript levels were normalized to loading control gene *eIF4A1*. One of the three biological replicates with a similar expression pattern was shown. Bars indicate the means ± SE of three technical replicates. **(B)** Genetic interaction between *35S:FIL* and *lfr-2*. The phenotype of whole plants (upper) and the representative rosette leaves (lower) from 22-day-old plants from different genotypes as indicated. Scale bars = 0.5 cm.

functional LFR-SWI3B complex, LFR might also regulate *IAMT1* expression independently of SWI3B. Taken together, our results demonstrate that LFR and SWI3B physically interact to directly regulate the expression of *FIL* and *IAMT1*. This provides an epigenetic mechanism underlying the development of flattened leaf lamina in Arabidopsis. The main difference between the current findings and our previous report (Lin et al., 2018) is that: our current findings revealed that the flattened leaf blade is regulated by LFR-SWI3B; our previous report showed that LFR interacts with AS2 to control the leaf petiole length, the formation of leaf midvein, and the elongation of leaflet-like structure at the leaf margin (Lin et al., 2018). Altogether, these explained the molecular mechanism underlying different aspects of Arabidopsis leaf development.

It was reported that RNAi mutants of *SWI3B* resulted in an upward-curling leaf phenotype resulting from decreased nucleosome occupation and increased transcript level of *IAMT1* (Han et al., 2018). However, whether *IAMT1* was the direct target of SWI3B and other targets of SWI3B and its interacting partners in leaf development remains largely unclear. In this study, we further identified LFR as the interacting partner of SWI3B in flattened leaf development and showed that *IAMT1* was a target of both LFR and SWI3B. In addition, we found that both proteins co-targeted *FIL*, which was a critical transcription factor involved in abaxial cell fate determination. Our previous work showed that LFR interacts with AS2 (Lin et al., 2018) and it was also demonstrated that the AS1-AS2 complex is functionally associated with the histone deacetylase HDA6 to regulate leaf development (Luo et al., 2012). Recently, it was reported that SWI3B interacts with HDA6 to maintain transposon silencing in Arabidopsis (Yang et al., 2020). And we found in this study that LFR interact with SWI3B and *SWI3B-amic* displayed a similar leaf margin phenotype (a sawtooth appearance) to *lfr* (**Figure 4**). All these results mutually supported that the ARM repeat domain-containing protein LFR might integrate the actions of transcription factors and epigenetic regulators into a concerted transcriptional complex to regulate the expression of some common target(s), such as *BP*. Furthermore, it was reported that the MONOPTEROS (MP/ARF5) transcription factor recruited the SWI/SNF chromatin remodelers, BRAHMA (BRM) and SPLAYED (SYD), to increase DNA accessibility of *FIL* for the induction of flower primordium initiation (Wu et al., 2015). In different tissues, both LFR-SWI3B and MP-BRM/SYD bind to similar regions of the *FIL* promoter (b and c loci). Therefore, the SWI/SNF chromatin remodeling complex members may play similar roles in regulating *FIL* expression in both leaf and flower (lateral organ) development. Therefore, it was interesting to test whether MP or other transcription factors recruited the LFR-SWI3B complex to the target genes.

Although *BRM* and *SYD* play similar positive regulatory roles on *FIL* in flower primordium initiation to that of LFR-SWI3B in leaf development, the *brm* and *syd* mutants displayed a downward-curling leaf phenotype (Sarnowski et al., 2005; Sacharowski et al., 2015), which was opposite to the phenotype of *LFR* loss-of-function and *SWI3B* knock-down mutants (**Figure 4**). The possible explanations are as follows: (1) the detailed tissue- or cell type-specific expression of *FIL* and other possible target genes may be different in different genotype backgrounds. (2) previous studies indicate that, in addition to similar functions, SWI/SNF subunits display distinct functions, such as those of SWI/SNF ATPase (e.g., BRM and SYD), SWI3 proteins (SWI3A, SWI3B, SWI3C, and SWI3D), and SWP73 (SWP73A and SWP73B) (Sarnowski et al., 2005; Bezhani et al., 2007; Sacharowski et al., 2015). In this study, we detected the physical interaction between LFR and SWI3B and possibly SWI3A. However, we did not detect any physical interaction between LFR and SWI3C and SWI3D (**Figure 4** and **Supplementary Figure 1**). These results suggest that LFR, SWI3B, and/or SWI3A may act in the same SWI/SNF chromatin remodeling complex in Arabidopsis leaf development. BRM is an ATPase subunit of the SWI/SNF complex. It physically interacts with SWI3C and SWP73B. A loss-of-function mutant exhibits a downward-curling leaf phenotype similar to that of *brm* (Hurtado et al., 2006; Sacharowski et al., 2015). Therefore, BRM, SWP73B, and SWI3C may be present in the same SWI/SNF complex in maintaining a flattened development process; LFR and SWI3B/3A may be present in another type of SWI/SNF complex, including another ATPase. (3) It is also possible that LFR-SWI3B may have functions independent of the SWISNF

complex. It is interesting to uncover the mechanisms underlying the differences in leaf phenotypes in these mutants. Whether LFR is a constant component of the SWI/SNF complex and the composition of different SWI/SNF complex in different tissues and developmental stages still need further investigation, which would shed light on the biochemical composition of SWI/SNF complex and the epigenetic control of plant development.

CONCLUSION

The results of our study indicate that LFR physically interacts with SWI3B, a core component of the SWI/SNF chromatin remodeling complex, and with the ND domain of LFR, which is responsible for the interaction between LFR and SWI3B. This interaction is crucial for LFR functions in Arabidopsis. Results of the genetic analysis further reveal that *lfr* and *SWI3B-amic* single and double mutants have upward-curling leaf phenotypes. This phenotype is similar to those with altered *FIL* and *IAMT1* expression. Moreover, the results of further experiments show that LFR binds to the chromatin of *FIL* and *IAMT1* and are partly dependent on SWI3B. And overexpression of *FIL* partly recovers the curly leaf defect of *lfr*. Taken together, LFR interacts with SWI3B to regulate *FIL* and *IAMT1* expression and maintains the normal leaf blade development process.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

SC and HZ planned and conceptualized the study and designed the experiments. XWL performed the BiFC, western blot assay, co-IP, ChIP-qPCR, partial phenotypic analysis, and qRT-PCR. CY did the Y2H, phenotypic analysis, and RT-PCR. BZ and SY performed the partial phenotypic analysis and transgenic plant screening. TY conducted the *35S:LFR-3FLAG/lfr-1* plasmid construction and partial phenotypic analysis. XRL performed plasmid construction. XWL, CY, HZ, and SC wrote the manuscript with input from co-authors. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (NSFC 31771351 and 31400240) and the Hebei Provincial Natural Science Foundation of China (C2020205025).

ACKNOWLEDGMENTS

We thank Tomasz J. Sarnowski (Laboratory of Plant Molecular Biology, Warsaw University) and Csaba Koncz (Max-Planck-Institut für Züchtungsforschung) for providing *swi3b-2*/+ seeds and anti-SWI3B antibodies. We thank Detlef Weigel (Max Planck Institute for Developmental Biology, Tübingen) for providing pRS300 plasmids. We thank Zhiyong Wang (Carnegie Institution, Stanford, CA, United States) for providing pMDC32 plasmids. Lastly, we also thank Ligeng Ma for providing the yeast two-hybrid cDNA library.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.717649/full#supplementary-material

Supplementary Figure 1 | Transgenic rescue lines of *35S:LFR-3FLAG/lfr-1*.

Supplementary Figure 2 | Transgenic rescue lines of *35S:SWI3B-3FLAG/3b-2*.

Supplementary Figure 3 | The interaction analysis between LFR and SWI3A, SWI3C, and SWI3D in yeast.

Supplementary Figure 4 | Self-activation analysis truncated LFR and SWI3B in yeast.

Supplementary Figure 5 | LFR and SWI3B did not associated with the chromatin of *YUC6*.

Supplementary Table 1 | Primers used in this paper.

Supplementary Table 2 | Potential LFR-interacting proteins identified by the Y2H screening.

REFERENCES

Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D., and Sonnhammer, E. L. (1999). Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res.* 27, 260–262. doi: 10.1093/nar/27.1.260

Bezhani, S., Winter, C., Hershman, S., Wagner, J. D., Kennedy, J. F., Kwon, C. S., et al. (2007). Unique, shared, and redundant roles for the Arabidopsis SWI/SNF chromatin remodeling ATPases BRAHMA and SPLAYED. *Plant Cell* 19, 403–416. doi: 10.1105/tpc.106.048272

Bonaccorso, O., Lee, J. E., Puah, L., Scutt, C. P., and Golz, J. F. (2012). Filamentous flower controls lateral organ development by acting as both an activator and a repressor. *BMC Plant Biol.* 12:176. doi: 10.1186/1471-2229-12-176

Bowman, J. L., Eshed, Y., and Baum, S. F. (2002). Establishment of polarity in angiosperm lateral organs. *Trends Genet.* 18, 134–141. doi: 10.1016/s0168-9525(01)02601-4

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

Du, F., Guan, C., and Jiao, Y. (2018). Molecular mechanisms of leaf morphogenesis. *Mol. Plant* 11, 1117–1134. doi: 10.1016/j.molp.2018.06.006

Efroni, I., Han, S. K., Kim, H. J., Wu, M. F., Steiner, E., Birnbaum, K. D., et al. (2013). Regulation of leaf maturation by chromatin-mediated modulation of cytokinin responses. *Dev. Cell* 24, 438–445. doi: 10.1016/j.devcel.2013.01.019

Gao, N., Wanf, Z.-J., Zeng, B., and Cui, S. (2008). Purification of Arabidopsis LFR recombinant protein in engineering bacteria and preparation of its antibody. *Prog. Biochem. Biophys.* 35, 1059–1064.

Han, W., Han, D., He, Z., Hu, H., Wu, Q., Zhang, J., et al. (2018). The SWI/SNF subunit SWI3B regulates IAMT1 expression via chromatin remodeling in Arabidopsis leaf development. *Plant Sci.* 271, 127–132. doi: 10.1016/j.plantsci.2018.03.021

Hargreaves, D. C., and Crabtree, G. R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res.* 21, 396–420. doi: 10.1038/cr.2011.32

Hurtado, L., Farrona, S., and Reyes, J. C. (2006). The putative SWI/SNF complex subunit BRAHMA activates flower homeotic genes in *Arabidopsis thaliana*. *Plant Mol. Biol.* 62, 291–304. doi: 10.1007/s11103-006-9021-2

Husbands, A. Y., Benkovics, A. H., Nogueira, F. T., Lodha, M., and Timmermans, M. C. (2015). The asymmetric leaves complex employs multiple modes of regulation to affect adaxial-abaxial patterning and leaf complexity. *Plant Cell* 27, 3321–3335. doi: 10.1105/tpc.15.00454

Jiang, J., Mao, N., Hu, H., Tang, J., Han, D., Liu, S., et al. (2019). A SWI/SNF subunit regulates chromosomal dissociation of structural maintenance complex 5 during DNA repair in plant cells. *Proc. Natl. Acad. Sci. U.S.A.* 116, 15288–15296. doi: 10.1073/pnas.1900308116

Jun, J. H., Ha, C. M., and Fletcher, J. C. (2010). Blade-on-petiole1 coordinates organ determinacy and axial polarity in arabidopsis by directly activating asymmetric leaves2. *Plant Cell* 22, 62–76. doi: 10.1105/tpc.109.070763

Lin, X., Gu, D., Zhao, H., Peng, Y., Zhang, G., Yuan, T., et al. (2018). LFR is functionally associated with AS2 to mediate leaf development in Arabidopsis. *Plant J.* 95, 598–612. doi: 10.1111/tpj.13973

Luo, M., Yu, C. W., Chen, F. F., Zhao, L., Tian, G., Liu, X., et al. (2012). Histone deacetylase HDA6 is functionally associated with AS1 in repression of KNOX genes in arabidopsis. *PLoS Genet.* 8:e1003114. doi: 10.1371/journal.pgen.1003114

Manuela, D., and Xu, M. (2020). Patterning a leaf by establishing polarities. *Front. Plant Sci.* 11:568730. doi: 10.3389/fpls.2020.568730

McConnell, J. R., and Barton, M. K. (1998). Leaf polarity and meristem formation in Arabidopsis. *Development* 125, 2935–2942. doi: 10.1242/dev.125.15.2935

Ou, B., Yin, K. Q., Liu, S. N., Yang, Y., Gu, T., Wing Hui, J. M., et al. (2011). A high-throughput screening system for Arabidopsis transcription factors and its application to Med25-dependent transcriptional regulation. *Mol. Plant* 4, 546–555. doi: 10.1093/mp/ssr002

Sacharowski, S. P., Gratkowska, D. M., Sarnowska, E. A., Kondrak, P., Jancewicz, I., Porri, A., et al. (2015). SWP73 subunits of Arabidopsis SWI/SNF chromatin remodeling complexes play distinct roles in leaf and flower development. *Plant Cell* 27, 1889–1906. doi: 10.1105/tpc.15.00233

Samuel, M. A., Salt, J. N., Shiu, S. H., and Goring, D. R. (2006). Multifunctional arm repeat domains in plants. *Int. Rev. Cytol.* 253, 1–26. doi: 10.1016/s0074-7696(06)53001-3

Sarnowski, T. J., Rios, G., Jasik, J., Swiezewski, S., Kaczanowski, S., Li, Y., et al. (2005). SWI3 subunits of putative SWI/SNF chromatin-remodeling complexes play distinct roles during Arabidopsis development. *Plant Cell* 17, 2454–2472. doi: 10.1105/tpc.105.031203

Sarnowski, T. J., Swiezewski, S., Pawlikowska, K., Kaczanowski, S., and Jerzmanowski, A. (2002). AtSWI3B, an Arabidopsis homolog of SWI3, a core subunit of yeast Swi/Snf chromatin remodeling complex, interacts with FCA, a regulator of flowering time. *Nucleic Acids Res.* 30, 3412–3421. doi: 10.1093/nar/gkf458

Sudarsanam, P., and Winston, F. (2000). The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. *Trends Genet.* 16, 345–351.

Sundaramoorthy, R., and Owen-Hughes, T. (2020). Chromatin remodelling comes into focus. *F1000Res.* 9:F1000FacultyRev–1011.

Vercruyssen, L., Verkest, A., Gonzalez, N., Heyndrickx, K. S., Eeckhout, D., Han, S. K., et al. (2014). Angustifolia3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during Arabidopsis leaf development. *Plant Cell* 26, 210–229. doi: 10.1105/tpc.113.115907

Vignali, M., Hassan, A. H., Neely, K. E., and Workman, J. L. (2000). ATP-dependent chromatin-remodeling complexes. *Mol. Cell Biol.* 20, 1899–1910.

Wang, X. T., Yuan, C., Yuan, T. T., and Cui, S. J. (2012). The Arabidopsis LFR gene is required for the formation of anther cell layers and normal expression of key regulatory genes. *Mol. Plant* 5, 993–1000. doi: 10.1093/mp/sss024

Wang, Z., Yuan, T., Yuan, C., Niu, Y., Sun, D., and Cui, S. (2009). LFR, which encodes a novel nuclear-localized Armadillo-repeat protein, affects multiple developmental processes in the aerial organs in Arabidopsis. *Plant Mol. Biol.* 69, 121–131. doi: 10.1007/s11103-008-9411-8

Wu, G., Lin, W. C., Huang, T., Poethig, R. S., Springer, P. S., and Kerstetter, R. A. (2008). Kanadi1 regulates adaxial-abaxial polarity in Arabidopsis by directly repressing the transcription of asymmetric leaves2. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16392–16397. doi: 10.1073/pnas.0803997105

Wu, M. F., Yamaguchi, N., Xiao, J., Bargmann, B., Estelle, M., Sang, Y., et al. (2015). Auxin-regulated chromatin switch directs acquisition of flower primordium founder fate. *Elife* 4:e09269.

Yamaguchi, N., Winter, C. M., Wu, M. F., Kwon, C. S., William, D. A., and Wagner, D. (2014). Protocols: chromatin immunoprecipitation from Arabidopsis tissues. *Arabidopsis Book* 12:e0170. doi: 10.1199/tab.0170

Yang, J., Yuan, L., Yen, M. R., Zheng, F., Ji, R., Peng, T., et al. (2020). SWI3B and HDA6 interact and are required for transposon silencing in Arabidopsis. *Plant J.* 102, 809–822. doi: 10.1111/tpj.14666

Yang, X., Zaurin, R., Beato, M., and Peterson, C. L. (2007). Swi3p controls SWI/SNF assembly and ATP-dependent H2A-H2B displacement. *Nat. Struct. Mol. Biol.* 14, 540–547. doi: 10.1038/nsmb1238

Yu, Y., Fu, W., Xu, J., Lei, Y., Song, X., Liang, Z., et al. (2021). Bromodomain-containing proteins BRD1, BRD2, and BRD13 are Core subunits of SWI/SNF complexes and are vital for their genomic targeting in Arabidopsis. *Mol. Plant* 14, 888–904. doi: 10.1016/j.molp.2021.03.018

Yu, Y., Liang, Z., Song, X., Fu, W., Xu, J., Lei, Y., et al. (2020). BRAHMA-interacting proteins BRIP1 and BRIP2 are core subunits of Arabidopsis SWI/SNF complexes. *Nat. Plants* 6, 996–1007. doi: 10.1038/s41477-020-0734-z

Yuan, C., Li, X.-R., Gu, D.-D., Gu, Y., Gao, Y.-J., and Cui, S.-J. (2012). The effect of Arabidopsis LFR protein domain on its co-transactivation and subcellular localization in nucleus. *Prog. Biochem. Biophys.* 39, 1003–1011. doi: 10.3724/sp.j.1206.2012.00048

Zhao, Z., Li, T., Peng, X., Wu, K., and Yang, S. (2019). Identification and characterization of tomato SWI3-like proteins: overexpression of SlSWIC increases the leaf size in transgenic Arabidopsis. *Int. J. Mol. Sci.* 20:5121. doi: 10.3390/ijms20205121

Zhu, Y., Rowley, M. J., Bohmdorfer, G., and Wierzbicki, A. T. (2013). A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Mol. Cell* 49, 298–309. doi: 10.1016/j.molcel.2012.11.011

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

REVIEW
published: 17 September 2021
doi: 10.3389/fpls.2021.726762

Small RNAs: The Essential Regulators in Plant Thermotolerance

*Zhi-Fang Zuo[1,2], Wenbo He[1], Jing Li[1], Beixin Mo[1] and Lin Liu[1]**

[1]*Guangdong Provincial Key Laboratory for Plant Epigenetics, College of Life Sciences and Oceanography, Longhua Bioindustry and Innovation Research Institute, Shenzhen University, Shenzhen, China,* [2]*Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering, Shenzhen University, Shenzhen, China*

Small RNAs (sRNAs) are a class of non-coding RNAs that consist of 21–24 nucleotides. They have been extensively investigated as critical regulators in a variety of biological processes in plants. sRNAs include two major classes: microRNAs (miRNAs) and small interfering RNAs (siRNAs), which differ in their biogenesis and functional pathways. Due to global warming, high-temperature stress has become one of the primary causes for crop loss worldwide. Recent studies have shown that sRNAs are involved in heat stress responses in plants and play essential roles in high-temperature acclimation. Genome-wide studies for heat-responsive sRNAs have been conducted in many plant species using high-throughput sequencing. The roles for these sRNAs in heat stress response were also unraveled subsequently in model plants and crops. Exploring how sRNAs regulate gene expression and their regulatory mechanisms will broaden our understanding of sRNAs in thermal stress responses of plant. Here, we highlight the roles of currently known miRNAs and siRNAs in heat stress responses and acclimation of plants. We also discuss the regulatory mechanisms of sRNAs and their targets that are responsive to heat stress, which will provide powerful molecular biological resources for engineering crops with improved thermotolerance.

Keywords: microRNA, small interfering RNA, heat stress response, regulatory mechanism, thermotolerance

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Bin Yu,
University of Nebraska-Lincoln, United States
Misato Ohtani,
University of Tokyo, Japan

***Correspondence:**
Lin Liu
linliu@szu.edu.cn

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

Received: *17 June 2021*
Accepted: *11 August 2021*
Published: *17 September 2021*

Citation:
Zuo Z-F, He W, Li J, Mo B and Liu L (2021) Small RNAs: The Essential Regulators in Plant Thermotolerance. Front. Plant Sci. 12:726762. doi: 10.3389/fpls.2021.726762

INTRODUCTION

As sessile organisms, plants are constantly exposed to a wide range of biotic and abiotic stresses that are unfavorable for their growth and development. Abiotic stresses, such as drought, salt, temperature, and heavy metals seriously impact the productivity of plants (Zhao et al., 2016; Zhu, 2016). Due to global warming, high-temperature stress has become one of the primary causes for crop loss (Liu et al., 2017a). Global yields of maize and wheat declined by 3.8 and 5.5%, respectively, due to temperature increases of approximately 0.13°C per decade since 1980 (Lobell et al., 2011). Consequently, the mechanisms for heat stress responses in plants have become a global concern and have received much attention.

Plants have evolved complex and diverse mechanisms to defend against ambient high-temperature stress and various factors are involved in plant thermotolerance, such as heat shock proteins (HSPs), reactive oxygen species (ROS)-scavenging enzymes, heat shock transcription factor (HSFs), and small RNAs (sRNAs; Ohama et al., 2017; Zhao et al., 2021). sRNAs are a class of non-coding RNAs that consist of 21–24 nucleotides (nt) and are critical regulators of

gene expression by causing either transcriptional gene silencing (TGS) or post-transcriptional gene silencing (PTGS; Axtell, 2013; D'Ario et al., 2017; Yu et al., 2019). Recently, sRNAs have been reported to participate in heat stress responses and play important roles in plant thermotolerance (Ruiz-Ferrer and Voinnet, 2009; Khraiwesh et al., 2012; Shriram et al., 2016; Liu et al., 2017a; Pagano et al., 2021). In this review, we focus on the roles and the regulatory mechanisms of sRNAs, mainly microRNAs (miRNAs) and small interfering RNAs (siRNAs) underlying heat stress tolerance in plants.

SMALL RNAs IN PLANTS

Endogenous sRNAs in plants are classified into two major types based on the tertiary subdivision by Axtell (2013): hairpin RNAs (hpRNAs) and siRNAs. Both hpRNAs and siRNAs result from cleavage of a double-stranded duplex from the helical region of larger RNA precursors by Dicer-like (DCL) enzyme (Axtell, 2013). hpRNAs are derived from a single-stranded RNA (ssRNA) precursor with a stem-loop hairpin structure, whereas siRNAs are derived from a double-stranded RNA (dsRNA) precursors. miRNAs are a well-studied subset of hpRNAs. siRNAs can be divided into two major subgroups including heterochromatic siRNAs (hc-siRNAs) and phased siRNAs (phasiRNAs; Axtell, 2013; Yu et al., 2019). Furthermore, trans-acting siRNAs (tasiRNAs) are a particular class of phasiRNAs that silence targets in trans. All these sRNAs differ in their biogenesis and modes of action.

MicroRNAs

The biogenesis and processing of miRNAs occur in multiple steps in plants (**Figure 1A**). (i) Similar to protein-coding genes, miRNA-encoded *MIR* genes are transcribed by RNA polymerase II (Pol II) to generate a long single-stranded primary miRNA (pri-miRNA), which is capped and polyadenylated in its 5' and 3' terminal regions, respectively (Lee et al., 2004; Xie et al., 2005). The pri-miRNA is predicted to form a stem-loop or hairpin secondary structure and the imperfectly paired double-stranded stem region contains the miRNA and miRNA* (Meyers et al., 2010). (ii) The pri-miRNA is first cleaved into a stem-loop miRNA precursor miRNA (pre-miRNA) and then the pre-miRNA is processed into miRNA-miRNA* duplex (Schauer et al., 2002; Kurihara and Watanabe, 2004). In this complex process, the RNase III enzyme DCL1 forms a nuclear dicing bodies (D-bodies) in the nucleus with two other partner proteins, HYPONASTIC LEAVES1 (HYL1) and SERRATE (SE), which ensure accurate and efficient splicing of pre-miRNA, resulting in the base of the stem being sliced (Yu et al., 2020). Other cofactors are involved for proper processing of pri-miRNA, such as the Cap-binding complex (CBC; Laubinger et al., 2008) and the Forkhead-associated (FHA) domain-containing protein Dawdle (DDL; Yu et al., 2008). The pre-miRNA without the base is then processed into a miRNA-miRNA* duplex by DCL1, which removes the hairpin loop. (iii) The miRNA-miRNA* duplex is further methylated by methyltransferase HUA ENHANCER1 (HEN1) to protect the 3' ends from uridylation (Li et al., 2005; Yu et al., 2005). (iv) In most cases, the methylated guide strand (miRNA) is incorporated into Argonaute-1 (AGO1) with the aid of TRANSPORIN 1 (TRN1); while the passenger strand (miRNA*) of the duplex is degraded. The AGO1-miRNA complex is exported from the nucleus to the cytoplasm *via* HASTY (HST; Bologna et al., 2018; Yu et al., 2019). (v) The mature miRNA associated with the miRNA-induced silencing complex (miRISC) is guided to target mRNAs for mRNA cleavage by AGO1 (transcript cleavage) or *via* inhibition of protein synthesis (translational repression; Llave et al., 2002; Iki et al., 2010; Li et al., 2013). In addition, the exonucleases SMALL RNA DEGRADING NUCLEASE1 (SDN1), nucleotidyl transferase HEN1 SUPPRESSOR1 (HESO1), and UTP: RNA URIDYLYLTRANSFERASE 1 (URT1) play critical roles in the process of miRNA turnover to regulate its steady-state level (Ramachandran and Chen, 2008; Zhao et al., 2012; Tu et al., 2015).

Small Interfering RNAs

Heterochromatic siRNAs

Heterochromatic siRNAs (hc-siRNAs) are derived from transposable and repetitive elements and are involved in DNA methylation or chromatin alteration by the canonical RNA-directed DNA methylation (RdDM) pathway (Matzke and Mosher, 2014; Borges and Martienssen, 2015). The biogenesis of hc-siRNAs requires RNA Polymerase IV (Pol IV) to transcribe the ssRNA, which is then synthesized into dsRNA by RDR2. The dsRNA is then processed into 24-nt siRNAs by DCL3 and methylated at their 3' ends by HEN1. After the siRNAs is incorporated into AGO4, RNA polymerase V (Pol V) transcribes transcripts that are complementary to the siRNA, and the siRNA-AGO4 complex is recruited. AGO4, RNA-DIRECTED DNA METHYLATION 1 (RDM1), and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) trigger *de novo* DNA methylation at the TGS level at symmetric CG and CHG sites, or asymmetric CHH sites (H stands for C, T, or A; **Figure 1B**; Matzke and Mosher, 2014; Du et al., 2015).

Phased siRNAs

Phased siRNAs (PhasiRNAs) originated from non-coding phasiRNA-generating (*PHAS*) loci in monocots or protein-coding genes in dicots, which are transcribed by Pol II (Deng et al., 2018; Yu et al., 2018). These target precursors are cleaved using a miRNA-mediated cleavage system: either "one-hit" (AGO1-miRNA complex) or "two-hit" (AGO7-miRNA complex) system. SUPPRESSOR OF GENE SILENCING3 (SGS3) stabilizes the primary siRNAs to protect them from AGO-mediated slicing (Yoshikawa et al., 2013). Then the primary siRNAs are converted into dsRNA by RDR6 and processed into 21- or 22-nt secondary siRNAs by DCL4 or DCL2, respectively. The 21-nt tasiRNAs, which originate from non-coding trans-acting siRNA (*TAS*) loci transcripts are recruited by the AGO1 complex to participate in the cleavage of target mRNAs (Adenot et al., 2006; Fukudome et al., 2011). The biogenesis of 21-nt phasiRNAs largely depends on DCL4 in rice; whereas a class of 24-nt phasiRNAs are processed by the DCL3 homolog DCL3b during the reproductive stage (Song et al., 2012; Komiya, 2017).

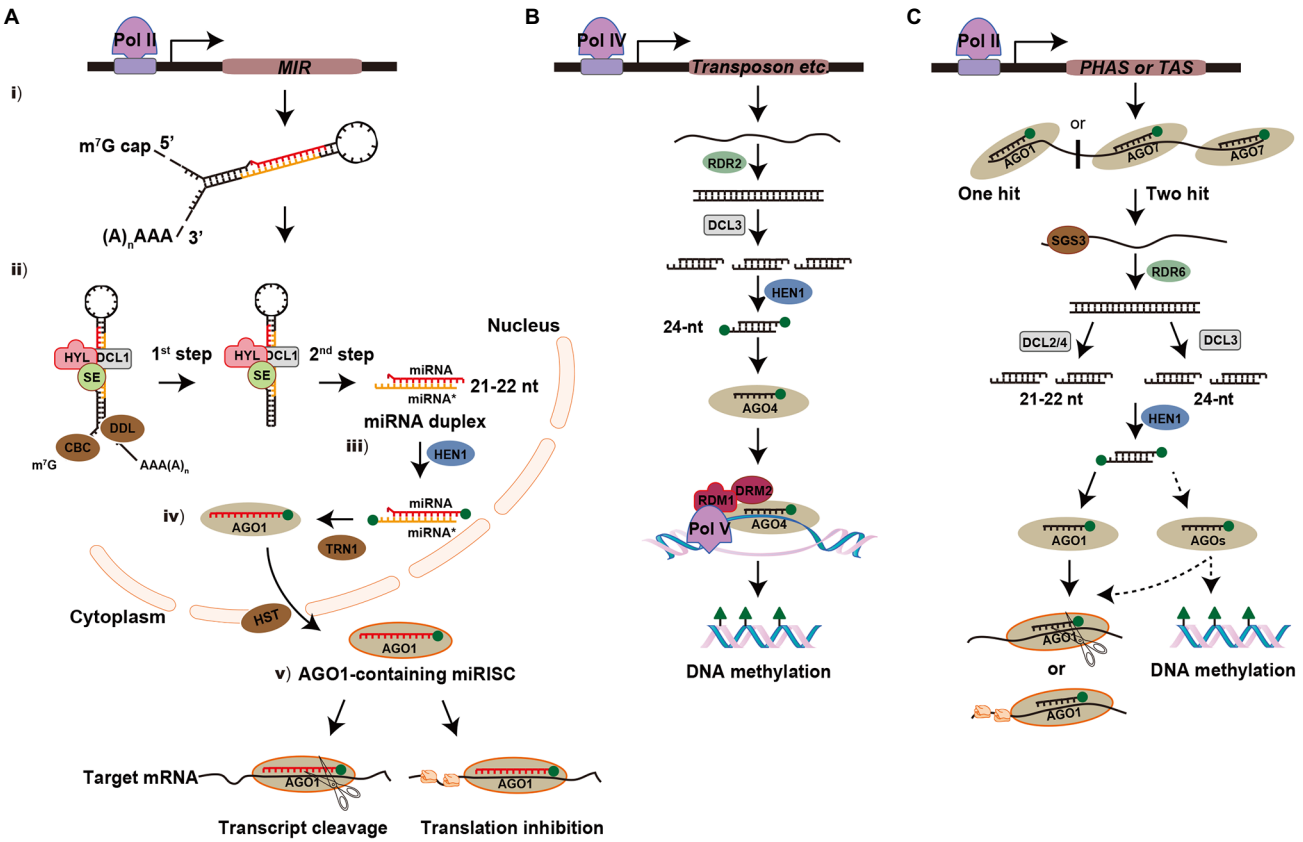

FIGURE 1 | Biogenesis and modes of action of microRNAs (miRNA) and small interfering RNAs (siRNA) in plants. **(A)** Overview of miRNA pathway. *MIR* gene is transcribed by polymerase II (Pol II) to generate a long single-stranded primary miRNA (pri-miRNA) and form a stem-loop or hairpin secondary structure. The pri-miRNA is firstly cleaved into stem-loop precursor miRNA (pre-miRNA) and then processed into miRNA-miRNA* duplex. In this complex process, DCL1 forms a D-bodies with HYPONASTIC LEAVES1 (HYL1) and SE to ensure the accurate and efficient splicing of pre-mRNA. Other cofactors are involved for proper processing, such as cap-binding complex (CBC) and Dawdle (DDL). The miRNA-miRNA* duplex is then methylated by HUA ENHANCER1 (HEN1) to protect the 3' ends from uridylation. The guide strand (miRNA) is incorporated into Argonaute-1 (AGO1) with the aid of TRANSPORIN 1 (TRN1) and then exported from the nucleus to the cytoplasm *via* HASTY (HST). In the cytoplasm, mature miRNA directs post-transcriptional gene silencing (PTGS) *via* transcript cleavage or translational repression. The green balls at the ends of miRNA-miRNA* duplex represent 2'-*O*-methy groups. **(B)** Overview of the heterochromatic siRNA (hc-siRNA) pathway. hc-siRNAs are derived from transposable and repetitive elements. The biogenesis of hc-siRNA requires Polymerase IV (Pol IV) to transcribe a single-stranded RNA (ssRNA) and synthesis into a double-stranded RNA (dsRNA) by RDR2. The dsRNA is then processed into 24-nt siRNA by DCL3. After the siRNA is incorporated into AGO4, Pol V transcribes transcripts that base pairing with siRNA and the siRNA-AGO4 complex is recruited. AGO4, RNA-DIRECTED DNA METHYLATION 1 (RDM1) and DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) trigger DNA methylation at the transcriptional gene silencing (TGS) level. **(C)** Overview of the phased siRNA (phasiRNA) pathway. phasiRNAs originated from phasiRNA-generating (*PHAS*) or trans-acting siRNA (*TAS*) loci are transcribed by Pol II and their precursors are cleaved *via* the miRNA "one-hit" system (AGO1-miRNA complex) or the "two-hit" system (AGO7-miRNA complex). SUPPRESSOR OF GENE SILENCING3 (SGS3) stabilizes the 5'- or 3'-cleaved fragments to protect them from AGO-mediated slicing. Then the primary siRNAs are converted into dsRNAs by RDR6 and further processed into 21- or 22-nt siRNAs by DCL2 and DCL4 or 24-nt siRNAs by DCL3. For further processing, some of them will be recruited by the AGO1 complex and some of them may be incorporated into other uncertain AGOs to trigger the cleavage of target mRNA, translational repression or DNA methylation.

However, little is known about the downstream processing of these phasiRNAs, such as which AGOs they interact with for silencing remains elusive (**Figure 1C**).

ROLES OF SMALL RNAs IN PLANT THERMOTOLERANCE

Global warming has diverse and profound effects on plant growth and development, and poses a serious threat to the global crop yields. Therefore, the plant response to high-temperature stress and the mechanism underlying plant thermotolerance have become focuses of research (Bita and Gerats, 2013). Recent studies have shown that plant miRNAs and siRNAs act as key regulators in response to high-temperature stress. Genome-wide studies for heat-responsive sRNAs have been conducted in many plant species using high-throughput sequencing and bioinformatics. A series of heat stress-responsive sRNAs have been identified from various plant species, suggesting that these sRNAs have persistent regulatory roles under extreme temperature (**Supplementary Table S1**).

miRNAs Involved in Heat Stress Responses

Extensive studies have shown that miRNAs can target genes encoding a diverse range of regulatory proteins, including a large proportion of TFs, suggesting that miRNAs function at the core of gene regulatory networks. One miRNA family usually has multiple target genes and plays versatile roles in several aspects of plant development and stress resistance. Accumulating evidence has shown that miRNAs are involved in plant responses to heat stress, and act as critical factors in coordinating plant development and heat stress resistance.

The miR156/miR172 Family

miR156 and its targets, *SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE* (*SPL*) genes are highly conserved in plants, and regulate developmental phase transitions, including juvenile-to-adult and vegetative-to-reproductive transitions (Xu et al., 2016; He et al., 2019; Ma et al., 2020). miR156 is highly expressed in young seedlings and its expression declines when the shoot develops, displaying opposite changes to its targets (Xu et al., 2016). The miR156-*SPL3* module regulates *flowering locus T* (*FT*) expression to control *Arabidopsis* flowering time in response to ambient temperature (16 and 23°C; Lee et al., 2010; Kim et al., 2012). Besides, miR156 was also found to response to high-temperature stress. Stief et al. (2014a) showed that miR156 isoforms were highly induced after recurring heat stress (37 and 44°C) and promoted sustained expression of heat stress responsive genes in *Arabidopsis*, suggesting miR156 was functionally important for heat stress memory. Heat stress memory refers to the maintenance of acquired thermotolerance that plants obtain after heat stress. Heat stress memory is one of the mechanisms for plants survival under recurring heat stress. Plants can withstand high temperature, which are lethal to them in a normal state. This acquired thermotolerance can be maintained several days after returning to non-stress temperatures (Lämke et al., 2016). miR156 mediated repression of *SPL2* and other target genes enhanced and prolonged the heat stress memory, and this process was also regulated by the HSFA2 cascade, which required HEAT STRESS-ASSOCIATED 32 (HSA32) and ROF1 (Stief et al., 2014a,b). In addition, recent studies showed that miR156-*SPL13* mediates heat stress response in alfalfa, and overexpression of soybean miR156b in *Arabidopsis* led to male sterility under heat stress (Matthews et al., 2019; Ding et al., 2021). These studies suggest that the role of the miR156-*SPL* module is conserved in plants and that plant development and heat stress responses are mediated by miRNAs.

In contrast to miR156, miR172 is a positive regulator in juvenile-to-adult developmental transition by targeting *APETALA2* (*AP2*) family genes, such as *AP2*, *TARGET OF EAT 1* (*TOE1*), *TOE2*, *TOE3*, *SCHLAFMUTZE* (*SMZ*), and *SCHNARCHZAPFEN* (*SNZ*; Gahlaut et al., 2018; Ma et al., 2020). However, miR156 is highly expressed in the juvenile phase; whereas miR172 is barely expressed in this stage (Wu et al., 2009). The targets of miR156, *SPL9*, and *SPL10*, directly promote the expression of miR172b, which indicates that miR172 acts downstream of miR156 to promote adult epidermal identity (Wu et al., 2009). miR172 has been reported to function in thermosensory pathway to regulate ambient temperature-responsive flowering under non-stress temperature conditions. The transgenic plants with overexpression of miR172 showed a temperature insensitive early flowering (Lee et al., 2010). Jung et al. (2012) discovered that RNA-binding protein FCA percept temperature fluctuation and promoted the processing of pri-miR172 *via* recognition of RNA motif in the stem-loop during the early stage of thermosensory flowering pathway. In addition, miR172 has also been revealed to response to high-temperature stress. In both rice post-meiosis panicle and safflower leaf tissues, miR172 was observed to be significantly downregulated, whereas its target *AP2* genes were upregulated under heat stress, indicating an important role of miR172-*AP2* module in plant heat stress response (Kouhi et al., 2020; Peng et al., 2020).

The miR159/miR319 Family

The miR159 and miR319 families are highly conserved in plants and have a high degree of sequence identity (Palatnik et al., 2019). miR159 targets several members of *GIBBERELLIC ACID MYB* (*GAMYB*) genes and plays important roles in flowering and male fertility. miR319 targets *TEOSINTE BRANCHED/CYCLOIDEA/PCF* (*TCP*) genes and several *MYB* genes to control leaf growth (Palatnik et al., 2019). miR159 was upregulated by heat stress in flowering Chinese cabbage (Ahmed et al., 2019). In contrast, miR159 was downregulated after heat stress in *Triticum aestivum*, and the transgenic rice overexpressing tae-miR159 showed sensitivity to heat stress (Wang et al., 2012). Heat stress caused a significant decrease of miR159 and an increase of its target genes, *CsGAMYB1* and *CsMYB29*-like in cucumber (Li et al., 2016). In addition, ectopic expression of csa-miR159b in *Arabidopsis* decreased heat tolerance by targeting *AtMYB33* (Li et al., 2016). For miR319, overexpression of sha-miR319d increased expression levels of heat stress-responsive genes and conferred heat stress tolerance in transgenic *Solanum lycopersicum* with increased activities of superoxide dismutase (SOD), catalase (CAT), and ascorbate peroxidase (APX; Shi et al., 2019).

The miR160/miR393 Family

The miR160 family targets the *AUXIN RESPONSIVE FACTOR* (*ARF*) gene family, which play vital roles in plant growth and development by regulating auxin signaling (Gahlaut et al., 2018). The miR160-*ARF* module is involved in phenotypic plasticity somatic embryo development, leaf development, root formation, and cell differentiation (Lin et al., 2015, 2018). In *Gossypium hirsutum*, overexpression of miR160 caused sensitivity to heat stress *via* suppressing the expression of *ARF10* and *ARF17* and caused anther indehiscence (Ding et al., 2017). Overexpression of a miR160 precursor presented increased thermotolerance, which shared similar phenotype with *arf10*, *arf16*, and *arf17* mutants in *Arabidopsis* (Lin et al., 2018). Furthermore, miR160 also regulates seed germination, hypocotyl, and rachis growth under heat stress (Lin et al., 2018).

As described above, miR160 is related to auxin. miR393 is also involved in auxin-related development in plants by

regulating the expression of the auxin receptors (TAARs) including TRANSPORT INHIBITOR RESPONSE1 (TIR1) and AUXIN SIGNALING F-BOX (AFBs). These TAARs can degrade the *AUXIN/INDOLE-3-ACETIC ACID* (*Aux/IAA*) genes and allow specific ARF TFs to active the auxin-responsive genes, which functions in the primary auxin-responsive pathway (Si-Ammour et al., 2011). Overexpression of osa-miR393a in transgenic creeping bentgrass increased heat tolerance by repressing its targets *AsAFB2* and *AsTIR1*, and this enhanced heat stress tolerance was associated with induced expression of *HSPs* (Zhao et al., 2019).

The miR398 Family

Heat stress causes the accumulation of ROS. SODs including iron SOD (Fe-SOD), manganese SOD (Mn-SOD), and copper/zinc SOD (Cu/Zn-SOD) encoded by *CSDs* are important ROS-scavenging enzymes that catalyze the superoxide radicals in plants. miR398 family members were found to be rapidly induced by heat stress, leading to the downregulation of their target genes *CSD1*, *CSD2*, and *Copper chaperone for SOD* (*CCS*) in *Arabidopsis* (Guan et al., 2013). Furthermore, *csd1*, *csd2*, and *ccs* mutants showed heat stress tolerance with increased expression of *HSFs* and *HSPs* (Guan et al., 2013). Fang et al. (2019) showed that the induction of tocopherols and 3'-phosphoadenosine 5'-phosphate (PAP) are required for the increased accumulation of miR398 and acquisition of heat tolerance. In addition, HSFA1b and HSFA7b were revealed to be responsible for the heat induction of miR398 by binding directly to the promoter of *MIR398* to activate its transcription (Guan et al., 2013). A recent study further revealed that the *MIR398* genes possess their natural antisense transcripts (NATs) and uncovered a regulatory loop between them; the *cis*-NATs of *MIR398* genes repress the processing of miR398 pri-miRNAs, which cause poorer thermotolerance due to the upregulation of miR398-targeted genes (Li et al., 2020). However, the underlying mechanism of how heat stress regulates the expression of *MIR398 cis*-NATs remains to be investigated.

Other miRNA Families

Other development-related miRNAs are also associated with plant heat stress responses. miR169 family members can target *Nuclear transcription factor Y subunit* (*NF-YA*) genes and function at flowering stage in rice under heat stress, which was confirmed by overexpression of miR169r-5p (Liu et al., 2017b). The paradigmatic miR396-*GROWTH-REGULATING FACTOR* (*GRF*) model is well established and plays important roles in regulating the size of multiple plant tissues or organs (Liu et al., 2021). Interestingly, Giacomelli et al. (2012) reported that miR396 mediated the cleavage of *HaWRKY6* in sunflower during early responses to high temperature. Heat stress reduced the accumulation of miR396, which showed opposite expression patterns to *HaWRKY6*, and expression of a miR396-resistant version of *HaWRKY6* altered heat stress responses in *Arabidopsis* (Giacomelli et al., 2012). For *MIR400* family, a heat stress-induced alternative splicing event was observed to occur in the intron of *MIR400*, which was co-transcribed with its host gene in *Arabidopsis* (Yan et al., 2012). Under heat stress, the alternative splicing of the *MIR400* intron resulted in greater accumulation of miR400 primary transcripts and reduced level of mature miR400. In addition, overexpression of miR400 caused higher sensitivity to heat stress in transgenic plants compare to the wild type plants. These results demonstrated that miR400 acts as a negative regulator in plant heat stress resistance and revealed the essential role of alternative splicing in linking miRNA and high-temperature stress (Yan et al., 2012). A report of rice miR5144-3p showed that miR5144-3p plays a role in protein folding and abiotic stress during rice development by regulating the expression of *OsPDIL1;1* (Xia et al., 2018). miR5144-3p was downregulated under heat stress, leading to the increased accumulation of *OsPDIL1;1* mRNA in rice; STTM-miR5144-3p and *OsPDIL1;1* overexpression transgenic rice exhibited improved heat tolerance (Xia et al., 2018).

siRNAs Involved in Heat Stress Responses

A large number of miRNAs and putative siRNAs participate in plant responses to environmental stresses, such as dehydration, salinity, cold, and ABA (Sunkar and Zhu, 2004). Compared with the large number of miRNAs and siRNAs identified in responses to other environmental stresses in plants, hitherto siRNAs identified in response to heat stress is relatively few. Previous researches have reported that heat-induced copia-type retrotransposon *ONSEN* was accumulated in the siRNAs biogenesis impaired mutants, which revealed the potential roles of siRNAs in plant heat stress response (Ito et al., 2011). The accumulation of a particular class of phasiRNAs-tasiRNAs derived from *Arabidopsis TAS* loci were found to decrease significantly under heat stress, indicating their participation in plant heat stress responses, and their functions in thermotolerance were subsequently investigated (Zhong et al., 2013; Li et al., 2014). Overexpression of *TAS1a*-derived tasiRNAs in *Arabidopsis* downregulated target genes *HEAT-INDUCED TAS1 TARGET1* (*HTT1*) and *HTT2*, and led to weaker thermotolerance in transgenic plants; whereas overexpression of *HTT1* and *HTT2* led to improved thermotolerance *via* upregulation of several *HSFs* (Li et al., 2014). Furthermore, heat-induced tasiRNA decrease was found to be involved in thermomemory of early flowering and attenuated immunity through targeting *HTT5*, which provides insights for understanding how heat exposure influence the fitness of plant progeny (Liu et al., 2019). Hu et al. (2020) have also identified a newly evolved phasiRNA locus that generated consecutive 21-nt phasiRNAs in response to heat stress in *Camellia*. Predictive bioinformatics and gene expression analysis showed that these secondary phasiRNAs could potentially target several genes including *LIPOXYGENASE*, *RAN GTPase*, *XYLOGLUCAN ENDOTRANSGLUCOSYLASE*, and *ATPase* to regulate their expression in a *trans*-acting manner. However, further genetic studies are required to verify these targets and elucidate the specific function of these phasiRNAs in thermal resistance. In a recent study, a large population of transposable element

derived 24-nt siRNAs were significantly reduced in maize tassels and roots after exposure to high temperature, and genes nearby these transposable elements tended to be downregulated, indicating that the expression of heat-dependent gene is influenced by adjacent transposon sequences. However, the underlying mechanism controlling the relationship among transposable element, transposable element derived 24-nt siRNAs and nearby genes in response to heat stress remains elusive (He et al., 2019).

The findings above demonstrate that sRNAs can serve as vital regulators of plant responses to heat stress. The major miRNAs and siRNAs involved in heat stress resistance and their regulatory pathway have been summarized in **Figure 2**. Interestingly, some sRNAs tend to display species-dependent expression patterns under heat stress, indicating that they may have distinct regulatory mechanisms in different plant species. For example, miR397 was downregulated in rice and tomato, but upregulated in banana in response to high temperature stress (Liu et al., 2017b; Pan et al., 2017; Zhu et al., 2019). In addition, several studies have identified novel species-specific sRNAs by genome-wide deep sequencing. For example, Liu et al. (2015b) identified 25 novel heat stress-responsive miRNAs in *Saccharina japonica*, such as sja-novel-mir-5, sja-novel-mir-13, and sja-novel-mir-59. These species-specific sRNAs may have essential roles in plant heat stress responses and dissecting their functions will broaden our understanding of the underlying regulatory mechanisms governing thermotolerance in different plant species.

MOLECULAR MECHANISMS OF PLANT RESPONSES TO HIGH-TEMPERATURE STRESS

Heat stress causes many adverse effects on plant growth, development, and physiological processes. Reduced water content, excess generation of ROS, and protein denaturation caused by heat stress greatly impede normal cellular functions (Hasanuzzaman et al., 2013; Jacob et al., 2017). Plants adapted to respond to and survive from heat stress by developing diverse mechanisms to cope with severe conditions, such as basal thermotolerance (Bokszczanin and Fragkostefanakis, 2013). Currently, heat stress response pathways have been extensively investigated in plants, and mainly include the HSP-based

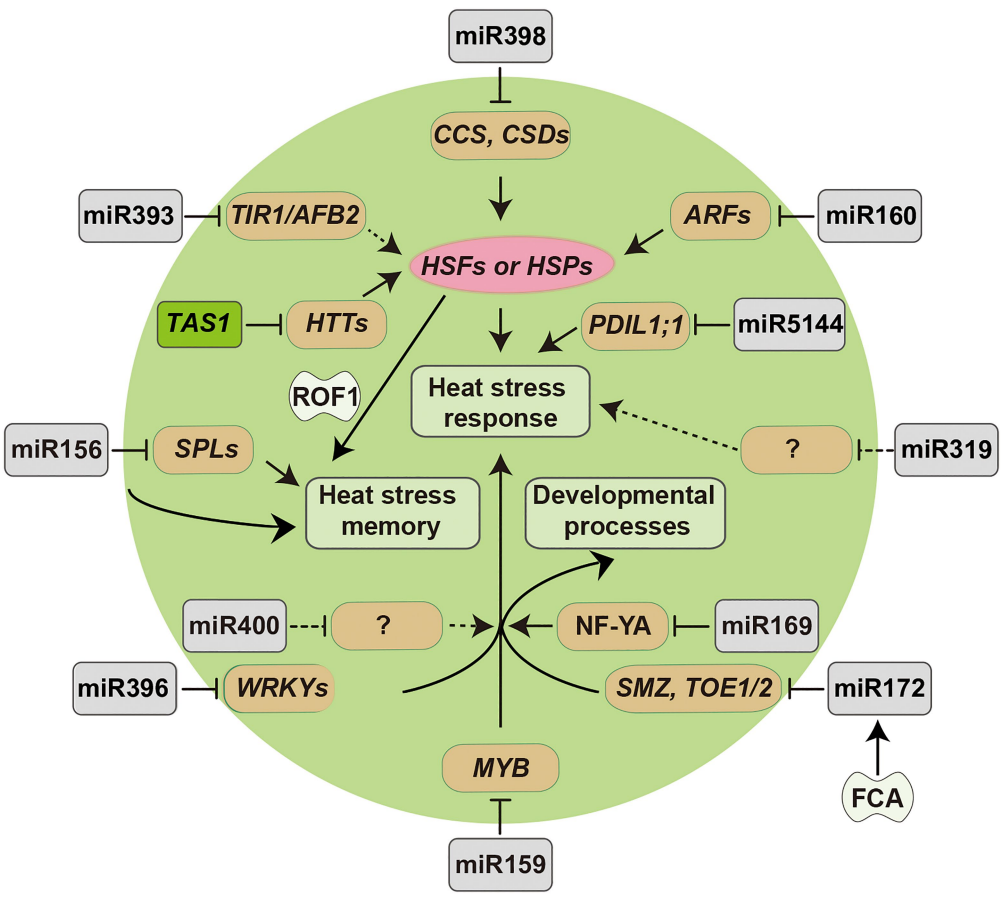

FIGURE 2 | Diagram of miRNA-target modules involved in plant responses to heat stress. Arrowed lines represent promotional effects and flat lines represent repressive effects. Dashed lines indicate hypotheses that need to be confirmed.

protective pathway, phospholipid pathway, ROS/redox signaling pathway, and phytohormone signaling pathway (Bokszczanin and Fragkostefanakis, 2013; Qu et al., 2013).

Heat-induced changes in plasma membrane fluidity cause a transient opening of Ca^{2+} channels, which induces Ca^{2+} influx to the cytoplasm. The increased levels of cytosolic Ca^{2+} activate multiple kinases, such as Calcium-dependent protein kinases (CDPKs), thereby evoking the expression of transcriptional regulators in response to heat stress. HSFs and HSPs play critical roles in this process (Qu et al., 2013; Guo et al., 2016). HSFs are evolutionarily grouped into A, B, and C classes. The *HSFA* subfamily genes, *HSFA1-HSFA9* have been well studied. *HSFA1s* function as master regulators, which are indispensable in acquired thermotolerance of plants by activating downstream heat stress responsive transcription factors (TFs), such as MULTIPROTEIN BRIDGING FACTOR1C (MBF1C) and DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN2A (DREB2A; Liu et al., 2011; Yoshida et al., 2011). HSPs, including HSP70, HSP90, HSP100, and HSP101 are the main inducers of HSFs, which resolve large aggregates induced by unfolded proteins and reduce protein misfolding (Jacob et al., 2017). In addition, HSP70 and HSP90 repress the activity of HSFA1s under non-stress condition by protein–protein interactions (Ohama et al., 2017). Phosphorylation and dephosphorylation by CALMODULIN-BINDING PROTEIN KINASE3 (CBK3) and PROTEIN PHOSPHATASE7 (PP7), respectively, can activate HSFA1s *via* post-translational modification under heat stress.

Heat stress causes remodeling of lipids in the membrane, which may induce the phospholipid signaling and the key mediators in this pathway include phosphatidyl inositol 4,5-bisphosphosphate (PIP2) and phosphatidic acid (PA). Accumulation of lipid signaling molecules can in turn trigger Ca^{2+} influx through channels in the plasma membrane (Bokszczanin and Fragkostefanakis, 2013). The mechanisms behind phospholipid signaling, from heat signal initiation to transduction, and the relationship between lipid signaling and plasma membrane channels as yet are still much unknown in plants.

Heat stress induces unfolded or misfolded proteins in the cytosol and endoplasmic reticulum (ER) that may trigger the unfolded protein response (UPR). Cytosolic UPR is mainly associated with specific HSFs, such as HSFA2, which is induced by HSFA1s and regulated by one splice variant of its own, S-HSfA2 *via* post-transcriptional regulation (Sugio et al., 2009; Liu and Charng, 2013). Two signaling pathways are involved in UPR, proteolytic processing, mediated by the basic leucine zipper domain (bZIP) TFs, RNA splicing, mediated by Inositol-requiring enzyme 1 (IRE1; Deng et al., 2013). ROS that accumulates during heat stress responses, such as H_2O_2, can act as signaling molecules to trigger the ROS-scavenging pathway; Respiratory burst oxidase homologs (RBOHs) also play important roles for the initiation and signal propagation of this pathway (Baxter et al., 2014). Furthermore, ROS-scavenging antioxidant enzymes, Late embryogenesis abundant (LEA) proteins, osmolytes, and secondary metabolites are considered necessary for detoxification of ROS (Bokszczanin and Fragkostefanakis, 2013).

Phytohormones are key players in plant growth and development from seed germination to senescence, and are involved in plant adaptation to adverse ambient stresses. Strigolactone (SL), cytokinin (CK), abscisic acid (ABA), and ethylene (Et) regulate the leaf senescence under heat stress (Abdelrahman et al., 2017). In addition, salicylic acid (SA) and jasmonic acid (JA), which are responsive to abiotic stresses, are also involved in the regulation of plant *HSF* genes (Guo et al., 2016; Rai et al., 2020). One phytohormone signaling pathway can respond to multiple abiotic stresses *via* crosstalk because plants endure multiple stresses in nature. For example, heat stress is usually associated with high light.

MECHANISMS OF SMALL RNA REGULATION IN PLANT THERMOTOLERANCE

Morphological Acclimation of Plants Under Heat Stress by sRNAs

Ambient temperature fluctuations affect plant functioning, geographical distribution, and agricultural production of crops (Proveniers and van Zanten, 2013). The phenotypic responses of plants to deal with high temperature include hypocotyl elongation, leaf hyponasty, and floral induction. The warm temperature induced basic helix–loop–helix (bHLH) TF, PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) plays a central role in warmth-mediated morphological acclimation (Franklin et al., 2011; Kumar et al., 2012; Kim et al., 2020). miRNAs are also involved in morphological adaptations of plant under heat stress. For example, overexpression of miR160 improved seed germination and increased the length of hypocotyl elongation and the rachis (Lin et al., 2018). The reproductive phase of flowering plants is highly sensitive to high temperature, which often contributes to the acceleration of flowering and results in poor seed set (Zinn et al., 2010). As indicated above, the heat-responsive miR156, miR159, miR172, and miR319 regulate flowering time or male and female fertility in different plant species, which illustrates that miRNAs play important roles in triggering the development of flower set (**Supplementary Table S1**; Wu et al., 2009; Lee et al., 2010; Stief et al., 2014a; Yin et al., 2018; Ahmed et al., 2019; Hu et al., 2019; Zhu et al., 2019; Kouhi et al., 2020). Furthermore, the heat-induced retrotransposon *ONSEN* was accumulated during flower development and before gametogenesis in mutants that are deficient in siRNA synthesis (Ito et al., 2011). Taken together, it is suggested that sRNAs play pivotal roles in regulating plant growth and reproductive tissue development under heat stress, which ensures transgenerational seed production (**Figure 3**).

Regulation of Essential Factors of HSFs/HSPs in Heat Stress Responses by sRNAs

Heat shock proteins that act as molecular chaperones are major functional proteins in heat stress response *via* the activation of HSFs (Qu et al., 2013; Ohama et al., 2017). Various attempts

have been made to increase thermotolerance in different host by overexpression of a single *HSF* or *HSP* gene, such as *CaHSP25.9*, *ZmHSF05*, *TaHSP23.9*, and *OsHSP20*, (Feng et al., 2019; Li et al., 2019; Guo et al., 2020; Wang et al., 2020). Nevertheless, *AsHSP26.8a*, a novel chloroplast-localized small *HSP* gene from creeping bentgrass negatively regulates heat stress resistance through modulating ABA and other stress signaling pathways (Sun et al., 2020). In addition, several miRNAs affect heat stress responses by targeting and activating *HSF/HSP* genes. The miR156-*SPL* module downregulates the expression of heat stress inducible genes, such as *HEAT STRESS ASSOCIATED 32* (*HSA32*), *HSP17.6A*, and *HSP22.0* during recovery from heat stress, which is functionally important for heat stress memory (Stief et al., 2014a). Overexpression of miR160 altered the expression of *HSPs* including *HSP17.6A*, *HSP17.6II*, *HSP21*, and *HSP70B*, which allow plants to survive under heat stress (Lin et al., 2018). In *Solanum habrochaites*, constitutive expression of sha-miR319d enhanced heat tolerance *via* upregulation of *HSFA1a* and *HSFA1b*, while overexpression of Osa-miR393a caused higher expression levels of *AsHSP17.0* and *AsHSP26.7a* than in wild type plants (Shi et al., 2019; Zhao et al., 2019). Transgenic plants expressing miR398-resistant forms of *CSD1*, *CSD2*, or *CCS* showed reduced expression levels of *HSF* genes (*HSFA1e*, *HSFA2*, *HSFA3*, and *HSFA7b*) and *HSP* genes (*HSP17.6*, *HSP70B*, and *HSP90.1*), while *csd1*, *csd2*, and *ccs* loss-of-function mutants showed enhanced thermotolerance with increased expression of heat stress inducible genes (Guan et al., 2013). Furthermore, HSFA1b and HSFA7b bind to heat stress elements directly in the promoter region of miR398b, which constituted a positive regulatory feedback loop (**Figure 3**). Overexpression of *HTT1* and *HTT2*, targets of *TAS1*, upregulated several *HSF* genes and enhanced thermotolerance in *Arabidopsis* (Li et al., 2014). Meanwhile, HSFA1a directly activates the expression of *HTT* genes, such as *HTT1*, which act as cofactors of HSP70-14 complexes in the thermotolerance pathway (Li et al., 2014). All of these

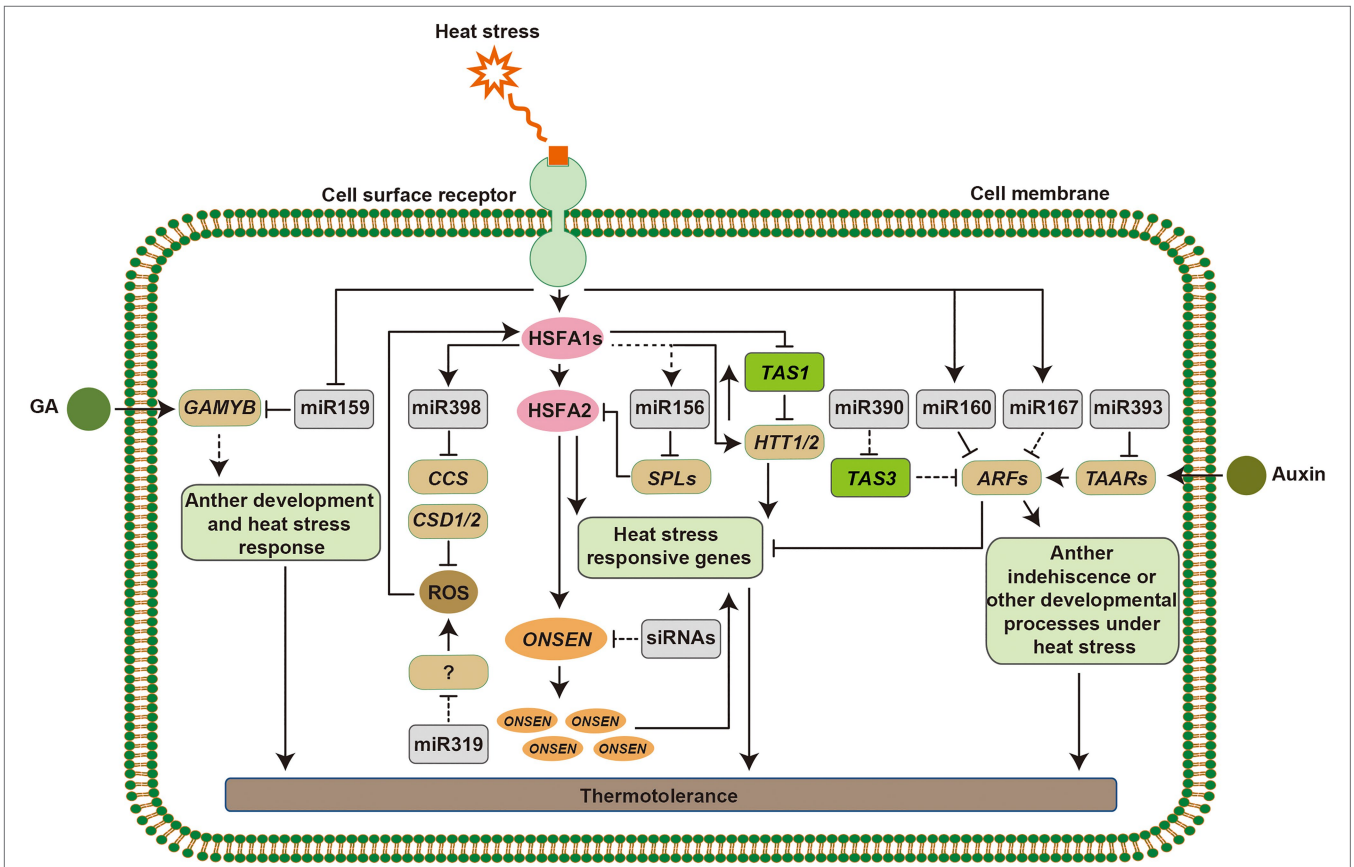

FIGURE 3 | Mechanism of small RNA (sRNA) regulation in plant responses to heat stress. Heat stress signaling can be perceived by putative sensors on the cell membrane, which induce the expression of heat stress responsive genes. HSFA1s function as master regulators to active the downstream heat stress responsive transcription factors and genes. Heat-induced miR398 downregulates *CSDs* and *CCS*. Accumulation of reactive oxygen species (ROS) in the cells acts as signals to alter the expression of heat shock transcription factors (*HSFs*), and miR319 functions in this pathway *via* unknown targets. HSFA1s directly activate the expression of *HTT* genes, targets of *TAS1*, which upregulate *HSF* and heat shock protein (*HSP*) genes. The miR156-*SPL* module promotes sustained expression of heat stress-inducible genes in response to recurring heat stress and prolong the heat stress memory. In addition, the retrotransposition *ONSEN* is activated by HSFAs, although, siRNAs regulate its activity, which can be inherited to its progeny. miR159-*GAMYB* modules function in gibberellin (GA) signaling. miR160, miR167, miR390, and miR393 target *AUXIN RESPONSIVE FACTOR* (*ARF*) or *TAARs* that act in the auxin response pathway and show morphological adaptations under heat stress. Arrowed lines represent promotional effects and flat lines represent repressive effects. Dashed lines indicate hypotheses that need to be confirmed.

discoveries reveal that sRNAs help plants achieve thermotolerance by regulating the expression of *HSFs* and *HSPs* (**Figure 3**).

sRNAs Mediated Heat Stress Responses Trigger the ROS-Scavenging Pathway

Unfavorable environments trigger the production of ROS, which causes oxidative damages to proteins, lipids and stress-induced electrolyte leakage in plants. Plants have evolved an antioxidant defense system equipped with various enzymatic and non-enzymatic ROS-scavengers to maintain an equilibrium between the production of ROS and elimination of excessive ROS (Caverzan et al., 2019). Several antioxidant enzymes, such as SOD, APX, and CAT are involved in the heat stress responses in the ROS-scavenging pathway, where oxidative stress is produced as a secondary stress (Qu et al., 2013). As described above, downregulation of *CSD1*, *CSD2*, and *CCS* by heat-induced miR398 led to accumulation ROS in cells, which contributed to the expression of *HSFs* and other heat stress-responsive genes (Guan et al., 2013). This regulatory mechanism constitutes a regulatory loop for plant thermotolerance that involves miR398, *HSF* genes, and ROS-scavenging enzymes. A recent study showed that, tocopherols and PAP positively regulated the biogenesis of miR398 in chloroplasts and promoted plant thermotolerance (Fang et al., 2019). In addition, altered ROS under heat stress may act as signals in cells to induce the expression of *HSF1Aa* and *HSF1Ab* in miR319d transgenic plants, but the mechanisms of ROS, miR319d, and its putative targets in this signal transduction are much unknown (Shi et al., 2019; **Figure 3**). These findings suggest that sRNAs-mediated heat stress responses are partially dependent on the ROS signaling pathways in plants.

Heat Stress Responsive sRNAs Involved in Phytohormone Signaling

Phytohormones are produced *via* environmental signals, such as heat stress. For example, miR159-regulated GAMYB-like TFs function in gibberellin (GA) signaling and overexpression of tae-miR159 caused sensitivity to heat stress in transgenic plants, which suggested that tae-miR159 may participate in a heat stress related signaling pathway (Murray et al., 2003; Wang et al., 2012). Auxin orchestrates many morphogenetic processes, such as root formation and anther development, and endogenous auxin is also involved in heat stress responses. miR160 and miR393 target *ARFs* or *TAARs*, which are components in the auxin response pathway (Si-Ammour et al., 2011; Kruszka et al., 2014; Gahlaut et al., 2018). Overexpression of miR160 in cotton increased sensitivity to heat stress and caused anther indehiscence by activating the auxin response; while constitutive expression of miR157 suppressed the auxin signal, which also caused sensitivity to heat stress with microspore abortion and anther indehiscence (Ding et al., 2017). Overexpressing of an miR160 precursor in *Arabidopsis* significantly reduced the expression of its target genes *ARF10*, *ARF16*, and *ARF17*, and led to improved tolerance of transgenic plants. Furthermore, *arf10*, *arf16*, and *arf17* mutants showed advanced thermotolerance by regulating the expression of *HSPs* (Lin et al., 2018). In addition, the miR167-*ARFs and* miR390-*TAS3*-*ARFs* models are involved in plants developmental processes; however, whether miR167 and miR390 function in heat stress response by targeting *ARFs* need to be tested (**Figure 3**). ABA induces the accumulation of miR168, and both plants overexpressing miR168a and loss-of-function mutant of its target *AGO1*, ago1-27, displayed ABA hypersensitivity and several abiotic stress tolerances (Li et al., 2012). miR168 also responds to heat stress in various species, such as *Arabidopsis*, rice, *Brassica rapa*, and flowering Chinese cabbage (Barciszewska-Pacak et al., 2015; Bilichak et al., 2015; Mangrauthia et al., 2017; Ahmed et al., 2019). Responses of plants to heat stress are complex, and may require physiological or metabolic changes from several phytohormone signaling pathways with a crosstalk, and there is no doubt that sRNAs are essential regulators in these processes.

sRNAs Are Involved in Heat Stress Memory

As discussed above, a multi-layered regulatory signaling pathways are involved in plant response to heat stress. Heat stress memory as one of the mechanisms for plants survival under recurring heat stress included the regulation of chromatin modifications. For example, high levels of histone H3 lysine 4 (H3K4) methylation are related to hyper-induction of heat stress inducible genes after a recurring heat stress, which depends on HSFA2 (Yoshida et al., 2011; Liu and Charng, 2013; Lämke et al., 2016). miR156 promotes sustained expression of heat stress inducible genes in response to recurring heat stress, and increases of miR156 prolong heat stress memory (Stief et al., 2014a,b). miR156 targets, *SPLs*, are critical for heat stress memory, and may serve to integrate morphological acclimation with heat stress responses. Furthermore, plants display transgenerational memory mediated by transposons. Transgenerational memory refers to transmitting epigenetic states or environmental responses from one generation to the next that may offer the offspring an adaptive advantage for better fitness (Liu et al., 2015a). A surprisingly high frequency of new heat-induced retrotransposition, *ONSEN* insertions in progeny after heat stress revealed that the transgenerational memory of heat stress is maintained during differentiation of generative organs by priming *ONSEN* to transpose (Ito et al., 2011). In addition, the activation of *ONSEN* requires heat-induced TFs in the heat stress response pathway, such as HSFA2 (Cavrak et al., 2014). Although, the mechanisms of DNA methylation, sRNAs, and transposons in heat stress memory are obscure, epigenetic regulation is an important mechanism in response to heat stress (**Figure 3**).

CONCLUSION AND PERSPECTIVES

Heat stress caused by the global warming affects the growth and development of plants, which increases the risk of yield reductions in agricultural crops. It is important to elucidate how plants respond to heat stress, but many questions remain to be answered, such as how do plants sense heat stress, and what kinds of signaling pathways the sensors use to transduce the signals into nucleus? Given that epigenetic regulation by

sRNAs is crucial in gene regulatory networks for heat stress responses, we believe that an intensive understanding of their roles and functions will provide plentiful potential biological resources for plant engineering. Recently, various heat stress responsive sRNAs have been identified. However, it should be noted that many experimental factors affect the responses of sRNAs in plants, such as plant species, tissue, developmental stage, treatment time, and growth conditions, which means that a slight change can lead to different expression patterns of sRNAs. We summarized the literatures on sRNAs involved in heat stress responses of plants and listed the heat stress conditions used for treatments in **Supplementary Table S1**. Even though, a large number of diverse heat stress responsive sRNAs have been identified in plants, their roles and molecular mechanisms are still not fully elucidated, which may be due to the lack of genetic materials, especially for non-model plants and important crops. Future research should focus on the creation of genetically modified mutants and genetic manipulations of sRNAs to identify additional sRNA-target modules in the regulatory networks of heat stress responses. Notably, most miRNAs that have been functionally investigated so far are evolutionary conserved miRNAs, whereas the species- or tissue-specific miRNAs have been rarely studied. Thus, much more work is required to decipher the regulatory mechanisms of non-conserved miRNAs in response to heat stress in more crops. In addition, the investigation of upstream regulation of heat-responsive miRNA would also be an interesting research topic and worthy of more attention.

As discussed above, several sRNAs were shown to have potential in improving plant thermotalerance. One type of sRNAs is involved in the heat stress responses by targeting and activating HSF/HSP genes that act as molecular chaperones to prevent denaturation or aggregation of target proteins, such as miR160 and miR393 (Lin et al., 2018; Zhao et al., 2019). Thus, increasing the saturation level of HSFs and HSPs by overexpression of one specific miRNA can enhance heat tolerance to their host plants using transgenic approaches. The other type of sRNAs participate in heat stress response by triggering ROS-scavenging pathway in plants, such as miR319 and miR398 (Guan et al., 2013; Shi et al., 2019). High levels of ROS-scavenging enzymes accumulation by manipulating certain miRNAs can be another strategy to generate heat stress tolerant plants. Recently, new biotechnological tools have been successfully explored to investigate *MIR* genes or miRNA modulation, such as endogenous artificial target mimicry (Short tandem target mimicry, STTM), miRNA transient virus induced gene silencing (VIGS) and *MIR* genes editing using CRISPR/Cas9 system (Basso et al., 2019). Furthermore, next-generation sequencing, for example, sRNA sequencing (sRNA-seq), Parallel Analysis of RNA Ends (PARE) analysis (Zhai et al., 2014; Jiang et al., 2020), and recently developed single-cell sRNA-mRNA co-sequencing (Wang et al., 2019) have provided powerful methods for elucidating the functions of sRNAs and their target genes. These techniques will be essential in further research and will expand the range of sRNA applications for crop breeding in thermotolerance.

Ultimately, the gene-silencing mechanisms mediated by sRNAs explore a new vista in the application of genetic engineering, which can be used to revolutionize agriculture by controlling a wide array of crop traits, including thermotolerance. sRNAs can work efficiently and precisely to develop targeted gene-silencing approaches in plants for various requirements, which can be used not only for the study of the functional analysis of genes responsive to heat stress but also to improve crop plants by manipulating their target genes. All of these accumulated researches will enable the successful and extensive application of sRNA technology for the development of next generation crops.

AUTHOR CONTRIBUTIONS

Z-FZ and LL wrote the manuscript. WH, JL, and BM provided the critical comments and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Shenzhen Grant Plan for Science and Technology (JCYJ20190808112207542), Guangdong Innovation Research Team Fund (2014ZT05S078), Natural Science Foundation of Guangdong Province (2019A1515011222 and 2021A1515010482), and Guangdong Basic and Applied Basic Research Foundation (2019A1515110162).

ACKNOWLEDGMENTS

We are thankful to reviewers for their valuable suggestions to improve this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.726762/full#supplementary-material

REFERENCES

Abdelrahman, M., El-Sayed, M., Jogaiah, S., Burritt, D. J., and Tran, L. P. (2017). The "STAY-GREEN" trait and phytohormone signaling networks in plants under heat stress. *Plant Cell Rep.* 36, 1009–1025. doi: 10.1007/s00299-017-2119-y

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouché, N., Gasciolli, V., et al. (2006). DRB4-dependent *TAS3 trans*-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* 16, 927–932. doi: 10.1016/j.cub.2006.03.035

Ahmed, W., Xia, Y., Zhang, H., Li, R., Bai, G., Siddique, K. H. M., et al. (2019). Identification of conserved and novel miRNAs responsive to heat stress in flowering Chinese cabbage using high-throughput sequencing. *Sci. Rep.* 9:14922. doi: 10.1038/s41598-019-51443-y

Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* 64, 137–159. doi: 10.1146/annurev-arplant-050312-120043

Ballén-Taborda, C., Plata, G., Ayling, S., Rodríguez-Zapata, F., Becerra Lopez-Lavalle, L. A., Duitama, J., et al. (2013). Identification of cassava microRNAs under abiotic stress. *Int. J. Genomics* 1–10. doi: 10.1155/2013/857986

Barciszewska-Pacak, M., Milanowska, K., Knop, K., Bielewicz, D., Nuc, P., Plewka, P., et al. (2015). *Arabidopsis* microRNA expression regulation in a wide range of abiotic stress responses. *Front. Plant Sci.* 6:410. doi: 10.3389/fpls.2015.00410

Basso, M. F., Ferreira, P. C. G., Kobayashi, A. K., Harmon, F. G., Nepomuceno, A. L., Molinari, H. B. C., et al. (2019). MicroRNAs and new biotechnological tools for its modulation and improving stress tolerance in plants. *Plant Biotechnol. J.* 17, 1482–1500. doi: 10.1111/pbi.13116

Baxter, A., Mittler, R., and Suzuki, N. (2014). ROS as key players in plant stress signalling. *J. Exp. Bot.* 65, 1229–1240. doi: 10.1093/jxb/ert375

Bilichak, A., Ilnytskyy, Y., Wóycicki, R., Kepeshchuk, N., Fogen, D., and Kovalchuk, I. (2015). The elucidation of stress memory inheritance in *Brassica rapa* plants. *Front. Plant Sci.* 6:5. doi: 10.3389/fpls.2015.00005

Bita, C. E., and Gerats, T. (2013). Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. *Front. Plant Sci.* 4:273. doi: 10.3389/fpls.2013.00273

Bokszczanin, K. L., and Fragkostefanakis, S. (2013). Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Front. Plant Sci.* 4:315. doi: 10.3389/fpls.2013.00315

Bologna, N. G., Iselin, R., Abriata, L. A., Sarazin, A., Pumplin, N., Jay, F., et al. (2018). Nucleo-cytosolic shuttling of ARGONAUTE1 prompts a revised model of the plant microRNA pathway. *Mol. Cell* 69, 709–719. doi: 10.1016/j.molcel.2018.01.007

Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085

Caverzan, A., Piasecki, C., Chavarria, G., Stewart, C. N., and Vargas, L. (2019). Defenses against ROS in crops and weeds: the effects of interference and herbicides. *Int. J. Mol. Sci.* 20:1086. doi: 10.3390/ijms20051086

Cavrak, V. V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L. M., and Mittelsten Scheid, O. (2014). How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genet.* 10:e1004115. doi: 10.1371/journal.pgen.1004115

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

D'Ario, M., Griffiths-Jones, S., and Kim, M. (2017). Small RNAs: big impact on plant development. *Trends Plant Sci.* 22, 1056–1068. doi: 10.1016/j.tplants.2017.09.009

Deng, P., Muhammad, S., Cao, M., and Wu, L. (2018). Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants. *Plant Biotechnol. J.* 16, 965–975. doi: 10.1111/pbi.12882

Deng, Y., Srivastava, R., and Howell, S. H. (2013). Endoplasmic reticulum (ER) stress response and its physiological roles in plants. *Int. J. Mol. Sci.* 14, 8188–8212. doi: 10.3390/ijms14048188

Ding, X., Guo, J., Zhang, Q., Yu, L., Zhao, T., and Yang, S. (2021). Heat-responsive miRNAs participate in the regulation of male fertility stability in soybean CMS-based F1 under high temperature stress. *Int. J. Mol. Sci.* 22:2446. doi: 10.3390/ijms22052446

Ding, Y., Ma, Y., Liu, N., Xu, J., Hu, Q., Li, Y., et al. (2017). microRNAs involved in auxin signalling modulate male sterility under high-temperature stress in cotton (*Gossypium hirsutum*). *Plant J.* 91, 977–994. doi: 10.1111/tpj.13620

Du, J., Johnson, L. M., Jacobsen, S. E., and Patel, D. J. (2015). DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* 16, 519–532. doi: 10.1038/nrm4043

Fang, X., Zhao, G., Zhang, S., Li, Y., Gu, H., Li, Y., et al. (2019). Chloroplast-to-nucleus signaling regulates microRNA biogenesis in *Arabidopsis*. *Dev. Cell* 48, 371.e4–382.e4. doi: 10.1016/j.devcel.2018.11.046

Feng, X. H., Zhang, H. X., Ali, M., Gai, W. X., Cheng, G. X., Yu, Q. H., et al. (2019). A small heat shock protein CaHsp25.9 positively regulates heat, salt, and drought stress tolerance in pepper (*Capsicum annuum* L.). *Plant Physiol. Biochem.* 142, 151–162. doi: 10.1016/j.plaphy.2019.07.001

Franklin, K. A., Lee, S. H., Patel, D., Kumar, S. V., Spartz, A. K., Gu, C., et al. (2011). Phytochrome-interacting factor 4 (PIF4) regulates auxin biosynthesis at high temperature. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20231–20235. doi: 10.1073/pnas.1110682108

Fukudome, A., Kanaya, A., Egami, M., Nakazawa, Y., Hiraguri, A., Moriyama, H., et al. (2011). Specific requirement of DRB4, a dsRNA-binding protein, for the in vitro dsRNA-cleaving activity of *Arabidopsis* Dicer-like 4. *RNA* 17, 750–760. doi: 10.1261/rna.2455411

Gahlaut, V., Baranwal, V. K., and Khurana, P. (2018). miRNomes involved in imparting thermotolerance to crop plants. *3 Biotech* 8:497. doi: 10.1007/s13205-018-1521-7

Giacomelli, J. I., Weigel, D., Chan, R. L., and Manavella, P. A. (2012). Role of recently evolved miRNA regulation of sunflower *HaWRKY6* in response to temperature damage. *New Phytol.* 195, 766–773. doi: 10.1111/j.1469-8137.2012.04259.x

Goswami, S., Kumar, R. R., and Rai, R. D. (2014). Heat-responsive microRNAs regulate the transcription factors and heat shock proteins in modulating thermo stability of starch biosynthesis enzymes in wheat (*Triticum aestivum* L.) under the heat stress. *Aust. J. Crop. Sci.* 8:697.

Guan, Q., Lu, X., Zeng, H., Zhang, Y., and Zhu, J. (2013). Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in *Arabidopsis*. *Plant J.* 74, 840–851. doi: 10.1111/tpj.12169

Guo, L. M., Li, J., He, J., Liu, H., and Zhang, H. M. (2020). A class I cytosolic HSP20 of rice enhances heat and salt tolerance in different organisms. *Sci. Rep.* 10:1383. doi: 10.1038/s41598-020-58395-8

Guo, M., Liu, J. H., Ma, X., Luo, D. X., Gong, Z. H., and Lu, M. H. (2016). The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Front. Plant Sci.* 7:114. doi: 10.3389/fpls.2016.00114

Hasanuzzaman, M., Nahar, K., Alam, M. M., Roychowdhury, R., and Fujita, M. (2013). Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *Int. J. Mol. Sci.* 14, 9643–9684. doi: 10.3390/ijms14059643

He, J., Jiang, Z., Gao, L., You, C., Ma, X., Wang, X., et al. (2019). Genome-wide transcript and small RNA profiling reveals transcriptomic responses to heat stress. *Plant Physiol.* 181, 609–629. doi: 10.1104/pp.19.00403

Hivrale, V., Zheng, Y., Puli, C. O. R., Jagadeeswaran, G., Gowdu, K., Kakani, V. G., et al. (2016). Characterization of drought- and heat-responsive microRNAs in switchgrass. *Plant Sci.* 242, 214–223. doi: 10.1016/j.plantsci.2015.07.018

Hu, Z., Lyu, T., Yan, C., Wang, Y., Ye, N., Fan, Z., et al. (2020). Identification of alternatively spliced gene isoforms and novel noncoding RNAs by single-molecule long-read sequencing in *camellia*. *RNA Biol.* 17, 966–976. doi: 10.1080/15476286.2020.1738703

Hu, Z., Shen, X., Xiang, X., and Cao, J. (2019). Evolution of MIR159/319 genes in *Brassica campestris* and their function in pollen development. *Plant Mol. Biol.* 101, 537–550. doi: 10.1007/s11103-019-00920-z

Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M. C., Matsumoto-Yokoyama, E., Mitsuhara, I., et al. (2010). In vitro assembly of plant RNA-induced silencing complexes facilitated by molecular chaperone HSP90. *Mol. Cell* 39, 282–291. doi: 10.1016/j.molcel.2010.05.014

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472, 115–119. doi: 10.1038/nature09861

Jacob, P., Hirt, H., and Bendahmane, A. (2017). The heat-shock protein/chaperone network and multiple stress resistance. *Plant Biotechnol. J.* 15, 405–414. doi: 10.1111/pbi.12659

Jiang, P., Lian, B., Liu, C., Fu, Z., Shen, Y., Cheng, Z., et al. (2020). 21-nt phasiRNAs direct target mRNA cleavage in rice male germ cells. *Nat. Commun.* 11, 1–10. doi: 10.1038/s41467-020-19034-y

Jodder, J., Das, R., Sarkar, D., Bhattacharjee, P., and Kundu, P. (2018). Distinct transcriptional and processing regulations control miR167a level in tomato during stress. *RNA Biol.* 15, 130–143. doi: 10.1080/15476286.2017.1391438

Jung, J. H., Seo, P. J., Ahn, J. H., and Park, C. M. (2012). *Arabidopsis* RNA-binding protein FCA regulates microRNA172 processing in thermosensory flowering. *J. Biol. Chem.* 287, 16007–16016. doi: 10.1074/jbc.M111.337485

Khraiwesh, B., Zhu, J.-K., and Zhu, J. (2012). Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta* 1819, 137–148. doi: 10.1016/j.bbagrm.2011.05.001

Kim, S., Hwang, G., Kim, S., Thi, T. N., Kim, H., Jeong, J., et al. (2020). The epidermis coordinates thermoresponsive growth through the phyB-PIF4-auxin pathway. *Nat. Commun.* 11:1053. doi: 10.1038/s41467-020-14905-w

Kim, J. J., Lee, J. H., Kim, W., Jung, H. S., Huijser, P., and Ahn, J. H. (2012). The *microRNA156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3* module regulates ambient temperature-responsive flowering via *FLOWERING*

LOCUS T in *Arabidopsis*. *Plant Physiol.* 159, 461–478. doi: 10.1104/pp.111.192369

Komiya, R. (2017). Biogenesis of diverse plant phasiRNAs involves an miRNA-trigger and Dicer-processing. *J. Plant Res.* 130, 17–23. doi: 10.1007/s10265-016-0878-0

Kouhi, F., Sorkheh, K., and Ercisli, S. (2020). MicroRNA expression patterns unveil differential expression of conserved miRNAs and target genes against abiotic stress in safflower. *PLoS One* 15:e0228850. doi: 10.1371/journal.pone.0228850

Kruszka, K., Pacak, A., Swida-Barteczka, A., Nuc, P., Alaba, S., Wroblewska, Z., et al. (2014). Transcriptionally and post-transcriptionally regulated microRNAs in heat stress response in barley. *J. Exp. Bot.* 65, 6123–6135. doi: 10.1093/jxb/eru353

Kumar, S. V., Lucyshyn, D., Jaeger, K. E., Alós, E., Alvey, E., Harberd, N. P., et al. (2012). Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature* 484, 242–245. doi: 10.1038/nature10928

Kumar, R. R., Pathak, H., Sharma, S. K., Kala, Y. K., Nirjal, M. K., Singh, G. P., et al. (2015). Novel and conserved heat-responsive microRNAs in wheat (*Triticum aestivum* L.). *Funct. Integr. Genomics* 15, 323–348. doi: 10.1007/s10142-014-0421-0

Kurihara, Y., and Watanabe, Y. (2004). *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12753–12758. doi: 10.1073/pnas.0403115101

Lämke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. *EMBO J.* 35, 162–175. doi: 10.15252/embj.201592593

Laubinger, S., Sachsenberg, T., Zeller, G., Busch, W., Lohmann, J. U., Rätsch, G., et al. (2008). Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8795–8800. doi: 10.1073/pnas.0802493105

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S. H., et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23, 4051–4060. doi: 10.1038/sj.emboj.7600385

Lee, H., Yoo, S. J., Lee, J. H., Kim, W., Yoo, S. K., Fitzgerald, H., et al. (2010). Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in *Arabidopsis*. *Nucleic Acids Res.* 38, 3081–3093. doi: 10.1093/nar/gkp1240

Li, W., Cui, X., Meng, Z., Huang, X., Xie, Q., Wu, H., et al. (2012). Transcriptional regulation of *Arabidopsis* MIR168a and argonaute1 homeostasis in abscisic acid and abiotic stress responses. *Plant Physiol.* 158, 1279–1292. doi: 10.1104/pp.111.188789

Li, H., Hu, T., Amombo, E., and Fu, J. (2017). Genome-wide identification of heat stress-responsive small RNAs in tall fescue (*Festuca arundinacea*) by high-throughput sequencing. *Plant Physiol.* 213, 157–165. doi: 10.1016/j.jplph.2017.03.004

Li, Y., Li, X., Yang, J., and He, Y. (2020). Natural antisense transcripts of *MIR398* genes suppress microR398 processing and attenuate plant thermotolerance. *Nat. Commun.* 11:5351. doi: 10.1038/s41467-020-19186-x

Li, S., Liu, J., Liu, Z., Li, X., Wu, F., and He, Y. (2014). *Heat-induced TAS1 TARGET1* mediates thermotolerance via heat stress transcription factor A1a-directed pathways in *Arabidopsis*. *Plant Cell* 26, 1764–1780. doi: 10.1105/tpc.114.124883

Li, S., Liu, L., Zhuang, X., Yu, Y., Liu, X., Cui, X., et al. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in *Arabidopsis*. *Cell* 153, 562–574. doi: 10.1016/j.cell.2013.04.005

Li, H., Wang, Y., Wang, Z., Guo, X., Wang, F., Xia, X. J., et al. (2016). Microarray and genetic analysis reveals that csa-miR159b plays a critical role in abscisic acid-mediated heat tolerance in grafted cucumber plants. *Plant Cell Environ.* 39, 1790–1804. doi: 10.1111/pce.12745

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in *Arabidopsis*. *Curr. Biol.* 15, 1501–1507. doi: 10.1016/j.cub.2005.07.029

Li, G. L., Zhang, H. N., Shao, H., Wang, G. Y., Zhang, Y. Y., Zhang, Y. J., et al. (2019). *ZmHsf05*, a new heat shock transcription factor from *Zea mays* L. improves thermotolerance in *Arabidopsis thaliana* and rescues thermotolerance defects of the *athsfa2* mutant. *Plant Sci.* 283, 375–384. doi: 10.1016/j.plantsci.2019.03.002

Lin, J. S., Kuo, C. C., Yang, I. C., Tsai, W. A., Shen, Y. H., Lin, C. C., et al. (2018). MicroRNA160 modulates plant development and heat shock protein gene expression to mediate heat tolerance in *Arabidopsis*. *Front. Plant Sci.* 9:68. doi: 10.3389/fpls.2018.00068

Lin, Y., Lai, Z., Tian, Q., Lin, L., Lai, R., Yang, M., et al. (2015). Endogenous target mimics down-regulate miR160 mediation of ARF10, −16, and −17 cleavage during somatic embryogenesis in Dimocarpus longan lour. *Front. Plant Sci.* 6:956. doi: 10.3389/fpls.2015.00956

Liu, H. C., and Charng, Y. Y. (2013). Common and distinct functions of *Arabidopsis* class A1 and A2 heat shock factors in diverse abiotic stress responses and development. *Plant Physiol.* 163, 276–290. doi: 10.1104/pp.113.221168

Liu, J., Feng, L., Gu, X., Deng, X., Qiu, Q., Li, Q., et al. (2019). An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in *Arabidopsis*. *Cell Res.* 29, 379–390. doi: 10.1038/s41422-019-0145-8

Liu, J., Feng, L., Li, J., and He, Z. (2015a). Genetic and epigenetic control of plant heat responses. *Front. Plant Sci.* 6:267. doi: 10.3389/fpls.2015.00267

Liu, H. C., Liao, H. T., and Charng, Y. Y. (2011). The role of class A1 heat shock factors (HSFA1s) in response to heat and other stresses in *Arabidopsis*. *Plant Cell Environ.* 34, 738–751. doi: 10.1111/j.1365-3040.2011.02278.x

Liu, F., Wang, W., Sun, X., Liang, Z., and Wang, F. (2015b). Conserved and novel heat stress-responsive microRNAs were identified by deep sequencing in *Saccharina japonica* (Laminariales, Phaeophyta). *Plant Cell Environ.* 38, 1357–1367. doi: 10.1111/pce.12484

Liu, Y., Yan, J., Wang, K., Li, D., Yang, R., Luo, H., et al. (2021). MiR396-GRF module associates with switchgrass biomass yield and feedstock quality. *Plant Biotechnol. J.* doi: 10.1111/pbi.13567 [Epub ahead of print]

Liu, Q., Yan, S., Yang, T., Zhang, S., Chen, Y., and Liu, B. (2017a). Small RNAs in regulating temperature stress response in plants. *J. Integr. Plant Biol.* 59, 774–791. doi: 10.1111/jipb.12571

Liu, Q., Yang, T. F., Yu, T., Zhang, S., Mao, X., Zhao, J., et al. (2017b). Integrating small RNA sequencing with QTL mapping for identification of miRNAs and their target genes associated with heat tolerance at the flowering stage in rice. *Front. Plant Sci.* 8:43. doi: 10.3389/fpls.2017.00043

Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002). Cleavage of scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297, 2053–2056. doi: 10.1126/science.1076311

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science* 333, 616–620. doi: 10.1126/science.1076311

Ma, J., Zhao, P., Liu, S., Yang, Q., and Guo, H. (2020). The control of developmental phase transitions by microRNAs and their targets in seed plants. *Int. J. Mol. Sci.* 21:1971. doi: 10.3390/ijms21061971

Mangrauthia, S. K., Bhogireddy, S., Agarwal, S., Prasanth, V. V., Voleti, S. R., Neelamraju, S., et al. (2017). Genome-wide changes in microRNA expression during short and prolonged heat stress and recovery in contrasting rice cultivars. *J. Exp. Bot.* 68, 2399–2412. doi: 10.1093/jxb/erx111

Matthews, C., Arshad, M., and Hannoufa, A. (2019). Alfalfa response to heat stress is modulated by microRNA156. *Physiol. Plant.* 165, 830–842. doi: 10.1111/ppl.12787

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408. doi: 10.1038/nrg3683

May, P., Liao, W., Wu, Y., Shuai, B., Mccombie, W. R., Zhang, M. Q., et al. (2013). The effects of carbon dioxide and temperature on microRNA expression in *Arabidopsis* development. *Nat. Commun.* 4:2145. doi: 10.1038/ncomms3145

Meyers, B. C., Simon, S. A., and Zhai, J. (2010). MicroRNA processing: battle of the bulge. *Curr. Biol.* 20, R68–R70. doi: 10.1016/j.cub.2009.12.008

Murray, F., Kalla, R., Jacobsen, J., and Gubler, F. (2003). A role for HvGAMYB in anther development. *Plant J.* 33, 481–491. doi: 10.1046/j.1365-313X.2003.01641.x

Ohama, N., Sato, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci.* 22, 53–65. doi: 10.1016/j.tplants.2016.08.015

Pagano, L., Rossi, R., Paesano, L., Marmiroli, N., and Marmiroli, M. (2021). miRNA regulation and stress adaptation in plants. *Environ. Exp. Bot.* 184:104369. doi: 10.1016/j.envexpbot.2020.104369

Palatnik, J. F., Wollmann, H., Schommer, C., Schwab, R., Boisbouvier, J., Rodriguez, R., et al. (2019). Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs miR159 and miR319. *Dev. Cell* 51:129. doi: 10.1016/j.devcel.2019.09.016

Pan, C., Ye, L., Zheng, Y., Wang, Y., Yang, D., Liu, X., et al. (2017). Identification and expression profiling of microRNAs involved in the stigma exsertion under high-temperature stress in tomato. *BMC Genomics* 18:843. doi: 10.1186/s12864-017-4238-9

Pandey, R., Joshi, G., Bhardwaj, A. R., Agarwal, M., and Katiyar-Agarwal, S. (2014). A comprehensive genome-wide study on tissue-specific and abiotic stress-specific mirnas in *Triticum aestivum*. *PLoS One* 9:e95800. doi: 10.1371/journal.pone.0095800

Peng, Y., Zhang, X., Liu, Y., and Chen, X. (2020). Exploring heat-response mechanisms of microRNAs based on microarray data of rice post-meiosis panicle. *Int. J. Genomics* 17:7582612. doi: 10.1155/2020/7582612

Proveniers, M. C., and Van Zanten, M. (2013). High temperature acclimation through PIF4 signaling. *Trends Plant Sci.* 18, 59–64. doi: 10.1016/j.tplants.2012.09.002

Qu, A. L., Ding, Y. F., Jiang, Q., and Zhu, C. (2013). Molecular mechanisms of the plant heat stress response. *Biochem. Biophys. Res. Commun.* 432, 203–207. doi: 10.1016/j.bbrc.2013.01.104

Ragupathy, R., Ravichandran, S., Mahdi, M. S., Huang, D., Reimer, E., Domaratzki, M., et al. (2016). Deep sequencing of wheat srna transcriptome reveals distinct temporal expression pattern of mirnas in response to heat, light and UV. *Sci. Rep.* 6:39373. doi: 10.1038/srep39373

Rai, K. K., Pandey, N., and Rai, S. P. (2020). Salicylic acid and nitric oxide signaling in plant heat stress. *Physiol. Plant.* 168, 241–255. doi: 10.1111/ppl.12958

Ramachandran, V., and Chen, X. (2008). Degradation of microRNAs by a family of exoribonucleases in *Arabidopsis*. *Science* 321, 1490–1492. doi: 10.1126/science.1163728

Ravichandran, S., Ragupathy, R., Edwards, T., Domaratzki, M., and Cloutier, S. (2019). MicroRNA-guided regulation of heat stress response in wheat. *BMC Genomics* 20:488. doi: 10.1186/s12864-019-5799-6

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

Sailaja, B., Voleti, S. R., Subrahmanyam, D., Sarla, N., Prasanth, V. V., Bhadana, V. P., et al. (2014). Prediction and expression analysis of miRNAs associated with heat stress in *Oryza sativa*. *Rice Sci.* 21, 3–12. doi: 10.1016/S1672-6308(13)60164-X

Schauer, S. E., Jacobsen, S. E., Meinke, D. W., and Ray, A. (2002). DICER-LIKE1: blind men and elephants in *Arabidopsis* development. *Trends Plant Sci.* 7, 487–491. doi: 10.1016/S1360-1385(02)02355-5

Shi, X., Jiang, F., Wen, J., and Wu, Z. (2019). Overexpression of *Solanum habrochaites* microRNA319d (sha-miR319d) confers chilling and heat stress tolerance in tomato (*S. lycopersicum*). *BMC Plant Biol.* 19:214. doi: 10.1186/s12870-019-1823-x

Shriram, V., Kumar, V., Devarumath, R. M., Khare, T. S., and Wani, S. H. (2016). MicroRNAs as potential targets for abiotic stress tolerance in plants. *Front. Plant Sci.* 7:817. doi: 10.3389/fpls.2016.00817

Si-Ammour, A., Windels, D., Arn-Bouldoires, E., Kutter, C., Ailhas, J., Meins, F., et al. (2011). miR393 and secondary siRNAs regulate expression of the *TIR1/AFB2* auxin receptor clade and auxin-related development of *Arabidopsis* leaves. *Plant Physiol.* 157, 683–691. doi: 10.1104/pp.111.180083

Song, X., Li, P., Zhai, J., Zhou, M., Ma, L., Liu, B., et al. (2012). Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. *Plant J.* 69, 462–474. doi: 10.1111/j.1365-313X.2011

Stief, A., Altmann, S., Hoffmann, K., Pant, B. D., Scheible, W. R., and Bäurle, I. (2014a). *Arabidopsis* miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* 26, 1792–1807. doi: 10.1105/tpc.114.123851

Stief, A., Brzezinka, K., Lämke, J., and Bäurle, I. (2014b). Epigenetic responses to heat stress at different time scales and the involvement of small RNAs. *Plant Signal. Behav.* 9:e970430. doi: 10.4161/15592316.2014.970430

Sugio, A., Dreos, R., Aparicio, F., and Maule, A. J. (2009). The cytosolic protein response as a subcomponent of the wider heat shock response in *Arabidopsis*. *Plant Cell* 21, 642–654. doi: 10.1105/tpc.108.062596

Sun, X., Zhu, J., Li, X., Li, Z., Han, L., and Luo, H. (2020). AsHSP26.8a, a creeping bentgrass small heat shock protein integrates different signaling pathways to modulate plant abiotic stress response. *BMC Plant Biol.* 20:184. doi: 10.1186/s12870-020-02369-5

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

Tu, B., Liu, L., Xu, C., Zhai, J., Li, S., Lopez, M. A., et al. (2015). Distinct and cooperative activities of HESO1 and URT1 nucleotidyl transferases in microRNA turnover in *Arabidopsis*. *PLoS Genet.* 11:e1005119. doi: 10.1371/journal.pgen.1005119

Wang, J., Gao, X., Dong, J., Tian, X., Wang, J., Palta, J. A., et al. (2020). Over-expression of the heat-responsive wheat gene *TaHSP23.9* in transgenic *Arabidopsis* conferred tolerance to heat and salt stress. *Front. Plant Sci.* 11:243. doi: 10.3389/fpls.2020.00243

Wang, Q., Liu, N., Yang, X., Tu, L., and Zhang, X. (2016). Small RNA-mediated responses to low- and high-temperature stresses in cotton. *Sci. Rep.* 6:35558. doi: 10.1038/srep35558

Wang, Y., Sun, F., Cao, H., Peng, H., Ni, Z., Sun, Q., et al. (2012). TamiR159 directed wheat *TaGAMYB* cleavage and its involvement in anther development and heat response. *PLoS One* 7:e48445. doi: 10.1371/journal.pone.0048445

Wang, N., Zheng, J., Chen, Z., Liu, Y., Dura, B., Kwak, M., et al. (2019). Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. *Nat. Commun.* 10:95. doi: 10.1038/s41467-018-07981-6

Wong, J., Gao, L., Yang, Y., Zhai, J., Arikit, S., Yu, Y., et al. (2014). Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. *Plant J.* 79, 928–940. doi: 10.1111/tpj.12590

Wu, G., Park, M. Y., Conway, S. R., Wang, J. W., Weigel, D., and Poethig, R. S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138, 750–759. doi: 10.1016/j.cell.2009.06.031

Xia, K., Zeng, X., Jiao, Z., Li, M., Xu, W., Nong, Q., et al. (2018). Formation of protein disulfide bonds catalyzed by OsPDIL1; 1 is mediated by MicroRNA5144-3p in rice. *Plant Cell Physiol.* 59, 331–342. doi: 10.1093/pcp/pcx189

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S. A., and Carrington, J. C. (2005). Expression of *Arabidopsis MIRNA* genes. *Plant Physiol.* 138, 2145–2154. doi: 10.1104/pp.105.062943

Xin, M., Wang, Y., Yao, Y., Xie, C., Peng, H., Ni, Z., et al. (2010). Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 10:123. doi: 10.1186/1471-2229-10-123

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

Yan, K., Liu, P., Wu, C. A., Yang, G. D., Xu, R., Guo, Q. H., et al. (2012). Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in *Arabidopsis thaliana*. *Mol. Cell* 48, 521–531. doi: 10.1016/j.molcel.2012.08.032

Yin, Z., Li, Y., Zhu, W., Fu, X., Han, X., Wang, J., et al. (2018). Identification, characterization, and expression patterns of *TCP* genes and microRNA319 in cotton. *Int. J. Mol. Sci.* 19:3655. doi: 10.3390/ijms19113655

Yoshida, T., Ohama, N., Nakajima, J., Kidokoro, S., Mizoi, J., Nakashima, K., et al. (2011). *Arabidopsis* HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Mol. Gen. Genomics.* 286, 321–332. doi: 10.1007/s00438-011-0647-7

Yoshikawa, M., Iki, T., Tsutsui, Y., Miyashita, K., Poethig, R. S., Habu, Y., et al. (2013). 3' fragment of miR173-programmed RISC-cleaved RNA is protected from degradation in a complex with RISC and SGS3. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4117–4122. doi: 10.1073/pnas.1217050110

Yu, B., Bi, L., Zheng, B., Ji, L., Chevalier, D., Agarwal, M., et al. (2008). The FHA domain proteins DAWDLE in *Arabidopsis* and SNIP1 in humans act in small RNA biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 105, 10073–10078. doi: 10.1073/pnas.0804218105

Yu, Y., Mo, X., and Mo, B. (2020). "Introduction to plant small RNAs," in *Plant Small RNA*. eds. P. Guleria and V. Kumar (London: Academic Press), 3–35.

Yu, X., Wang, H., Lu, Y., De Ruiter, M., Cariaso, M., Prins, M., et al. (2012). Identification of conserved and novel microRNAs that are responsive to heat stress in brassica rapa. *J. Exp. Bot.* 63, 1025–1038. doi: 10.1093/jxb/err337

Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., et al. (2005). Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935. doi: 10.1126/science.1107130

Yu, Y., Zhang, Y., Chen, X., and Chen, Y. (2019). Plant noncoding RNAs: hidden players in development and stress responses. *Annu. Rev. Cell Dev. Biol.* 35, 407–431. doi: 10.1146/annurev-cellbio-100818-125218

Yu, Y., Zhou, Y., Zhang, Y., and Chen, Y. (2018). Grass phasiRNAs and male fertility. *Sci. China Life Sci.* 61, 148–154. doi: 10.1007/s11427-017-9166-3

Zhai, J., Arikit, S., Simon, S. A., Kingham, B. F., and Meyers, B. C. (2014). Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. *Methods* 67, 84–90. doi: 10.1016/j.ymeth.2013.06.025

Zhao, J., He, Q., Chen, G., Wang, L., and Jin, B. (2016). Regulation of non-coding RNAs in heat stress responses of plants. *Front. Plant Sci.* 7:1213. doi: 10.3389/fpls.2016.01213

Zhao, J., Lu, Z., Wang, L., and Jin, B. (2021). Plant responses to heat stress: physiology, transcription, noncoding RNAs, and epigenetics. *Int. J. Mol. Sci.* 22:117. doi: 10.3390/ijms22010117

Zhao, Y., Yu, Y., Zhai, J., Ramachandran, V., Dinh, T. T., Meyers, B. C., et al. (2012). The *Arabidopsis* nucleotidyl transferase HESO1 uridylates unmethylated small RNAs to trigger their degradation. *Curr. Biol.* 22, 689–694. doi: 10.1016/j.cub.2012.02.051

Zhao, J., Yuan, S., Zhou, M., Yuan, N., Li, Z., Hu, Q., et al. (2019). Transgenic creeping bentgrass overexpressing Osa-miR393a exhibits altered plant development and improved multiple stress tolerance. *Plant Biotechnol. J.* 17, 233–251. doi: 10.1111/pbi.12960

Zhong, S. H., Liu, J. Z., Jin, H., Lin, L., Li, Q., Chen, Y., et al. (2013). Warm temperatures induce transgenerational epigenetic release of RNA silencing by inhibiting siRNA biogenesis in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 110, 9171–9176. doi: 10.1073/pnas.1219655110

Zhou, R., Wang, Q., Jiang, F., Cao, X., Sun, M., Liu, M., et al. (2016). Identification of miRNAs and their targets in wild tomato at moderately and acutely elevated temperatures by high-throughput sequencing and degradome analysis. *Sci. Rep.* 6:33777. doi: 10.1038/srep33777

Zhu, J.-K. (2016). Abiotic stress signaling and responses in plants. *Cell* 167, 313–324. doi: 10.1016/j.cell.2016.08.029

Zhu, H., Zhang, Y., Tang, R., Qu, H., Duan, X., and Jiang, Y. (2019). Banana sRNAome and degradome identify microRNAs functioning in differential responses to temperature stress. *BMC Genomics* 20:33. doi: 10.1186/s12864-018-5395-1

Zinn, K. E., Tunc-Ozdemir, M., and Harper, J. F. (2010). Temperature stress and plant sexual reproduction: uncovering the weakest links. *J. Exp. Bot.* 61, 1959–1968. doi: 10.1093/jxb/erq053

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: 30 November 2021
doi: 10.3389/fpls.2021.765383

Dynamic Changes of DNA Methylation During Wild Strawberry (*Fragaria nilgerrensis*) Tissue Culture

Qiang Cao[1†], Yuxi Feng[1†], Xiongwei Dai[1†], Lin Huang[1], Jiamin Li[1], Pang Tao[2], M. James C. Crabbe[3,4,5], Ticao Zhang[6] and Qin Qiao[1]**

[1] School of Agriculture, Yunnan University, Kunming, China, [2] Horticultural Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, China, [3] Wolfson College, Oxford University, Oxford, United Kingdom, [4] Institute of Biomedical and Environmental Science and Technology, School of Life Sciences, University of Bedfordshire, Luton, United Kingdom, [5] School of Life Sciences, Shanxi University, Taiyuan, China, [6] College of Chinese Material Medica, Yunnan University of Chinese Medicine, Kunming, China

OPEN ACCESS

Edited by:
Gang Wu,
Zhejiang Agriculture and Forestry University, China

Reviewed by:
Chunying Kang,
Huazhong Agricultural University, China
Yuntao Zhang,
Beijing Academy of Agricultural and Forestry Sciences, China

***Correspondence:**
Ticao Zhang
ticaozhang@126.com
Qin Qiao
qiaoqin@ynu.edu.cn

[†] These authors have contributed equally to this work

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

Received: *27 August 2021*
Accepted: *09 November 2021*
Published: *30 November 2021*

Citation:
Cao Q, Feng Y, Dai X, Huang L, Li J, Tao P, Crabbe MJC, Zhang T and Qiao Q (2021) Dynamic Changes of DNA Methylation During Wild Strawberry (Fragaria nilgerrensis) Tissue Culture. Front. Plant Sci. 12:765383. doi: 10.3389/fpls.2021.765383

Tissue culture is an important tool for asexual propagation and genetic transformation of strawberry plants. In plant tissue culture, variation of DNA methylation is a potential source of phenotypic variation in regenerated plants. However, the genome wide dynamic methylation patterns of strawberry tissue culture remain unclear. In this study, we used whole-genome bisulfite sequencing (WGBS) to study genomic DNA methylation changes of a wild strawberry *Fragaria nilgerrensis* at six stages: from explants of shoot tips to outplanting and acclimation. Global methylation levels showed that CG sites exhibited the highest methylation level in all stages with an average of 49.5%, followed by CHG (33.2%) and CHH (12.4%). Although CHH accounted for the lowest proportion of total cytosine methylation, it showed the most obvious methylation change and the most of these changes occurred in the transposable element regions. The overall methylation levels alternately decreased and increased during the entire tissue culture process and the distribution of DNA methylation was non-uniform among different genetic regions. Furthermore, much more differentially methylated regions (DMRs) were detected in dedifferentiation and redifferentiation stages and most of them were transposable elements, suggesting these processes involved activating or silencing of amounts of transposons. The functional enrichment of the DMR-related genes indicated that genes involved in hormone metabolic processes, plant development and the stress response changed methylation throughout the tissue culture process. Finally, the quantitative real-time PCR (qRT-PCR) was conducted to examine the association of methylation and gene expression of a set of different methylated genes. Our findings give deeper insight into the epigenetic regulation of gene expression during the plant tissue cultures process, which will be useful in the efficient control of somaclonal variations and in crop improvement.

Keywords: ***Fragaria nilgerrensis*, tissue culture, somaclonal variations, DNA methylation, gene expression**

INTRODUCTION

The strawberry is one of the most economically important fruits in the world, belonging to the genus *Fragaria* L. (Rosaceae). *Fragaria nilgerrensis* is a widely distributed diploid wild strawberry in southwest China. Its white fruits with a unique peach aroma, as well as strong resistance to drought and cold are valuable characteristics for cultivated strawberry improvement (Noguchi et al., 2002; Guo et al., 2018; Lu et al., 2021). Recently, the genome sequence of *F. nilgerrensis* has been released and it could serve as another ideal model system for genetic studies of strawberry plants, and has great potential in broadening the genetic background of cultivated strawberries (Feng et al., 2020; Qiao et al., 2021).

The plant tissue culture technique is one of the most important tools in modern plant science research, which can be used for rapid asexual reproduction and genetic transformation, as well as an important means to understand the cell totipotency of plants (Ghosh et al., 2021). Under the influence of artificial hormonal environments, plant cells need to reset their genetic and epigenetic programs to adapt to the *in vitro* culture environment, and such molecular dynamic changes can also lead to stable genetic or epigenetic variations in clone progeny, also known as "somatic variation." These mutations may not be conductive to commercial production from tissue culture, but they are an important source for the development of new varieties with particular characteristics.

Among epigenetic factors, DNA methylation plays an important role in regulating chromatin conformation and gene expression during plant regeneration (Gupta et al., 2006; Ehrlich and Lacey, 2013; Lee and Seo, 2018). It has been reported that alteration of DNA methylation is related to developmental switches occurring during *in vitro* culture, which is determined by several factors including plant growth regulators, genetic backgrounds, and different types of stress (Baránek et al., 2010; Us-Camas et al., 2014; Karim et al., 2016). Recent advances in the field of epigenetics have revealed highly dynamic mechanisms of global and local DNA methylation variations occurring during cell dedifferentiation and redifferentiation processes in callus formation (Horstman et al., 2017; Xia et al., 2017). Few studies have focused on dynamic changes of methylation patterns during the whole process of tissue culture, which is not only crucial for commercial production of disease-free strawberry plants, but also for constructing a genetic transformation system. Understanding the epigenetic landscape and epigenetic mechanisms that modulate gene expression at each stage of tissue culture may be crucial for understanding variant phenotypes. This information can be used in crop improvement programs in a controlled way to generate better agronomic traits based on selection for favorable epigenetic states, creation of novel epialleles and avoided the negative consequences of variation (Tetsu and Akemi, 2013).

Therefore, in the present study, we explored the genome-wide methylation patterns and differences at the CG, CHG and CHH sites of six developmental stages of tissue culture in *F. nilgerrensis*. The differentially methylated regions (DMRs) were detected between each adjacent stage and associated genes with altered methylation were identified. Our results will help to identify the hypervariable regions in the plant genome during the tissue culture process, which should lead to the efficient control of somaclonal variations and their use in crop improvement programs.

MATERIALS AND METHODS

Plant Material and Tissue Culture

Plants of *F. nilgerrensis* were grown in the greenhouse in Yunnan University and conventionally propagated by runners to ensure all the plant materials were from the same clone. Runner tips 1–2 cm long were taken from these plants as explants. Explants were rinsed under running tap water for 30 min and then immersed in 75% alcohol for 20–25 s, followed by 0.1% $HgCl_2$ for 7 min. After that, the explants were thoroughly washed (4–5 washings) with sterilized distilled water and then shortened to 3–5 mm long. Finally, they were sampled or transferred to optimized medium for strawberry micropropagation in turn as listed in **Table 1**. The tissue culture was conducted in an incubation room at 14/10 light/day photoperiod conditions (38 $\mu E\ m^{-2}\ s^{-1}$) at temperatures of 25 $\pm$ 2°C for day and 20 $\pm$ 2°C for night. In the callus induction stage, dark culture lasting about 10 days was first required. The tissue cultured plantlets were transferred to pots in the greenhouse after proper hardening. The culture medium used for each stage is shown in **Table 1**. The materials collected from each stage with three biological replicates were shock-frozen in liquid nitrogen immediately and stored at −80°C.

Library Construction and Whole-Genome Bisulfite Sequencing

Genomic DNA was extracted using the Hi-DNAsecure Plant Kit (Qiagen GmbH, Hilden, Germany), according to the manufacturer's recommendations. Genomic DNA degradation and contamination was monitored on agarose gels. A total of 5.2 ug qualified genomic DNA spiked with 26 ng lambda DNA was fragmented by sonication to 200–300 bp with a Covaris S220 (Covaris, Woburn, MA, United States), followed by end repair and adenylation. Cytosine-methylated barcodes were ligated to sonicated DNA as per the manufacturer's instructions. Then these DNA fragments were treated twice with bisulfite using the EZ DNA Methylation-Gold^{TM} Kit (Zymo Research, Irvine, CA, United States). The resulting single-strand DNA fragments were amplified by the polymerase chain reaction (PCR) using KAPA HiFi HotStart Uracil + ReadyMix (2X) (Kapa Biosystems, Wilmington, MA, United States). Library concentration was determined with a Qubit 2.0 Flurometer (Life Technologies, CA, United States) and quantitative PCR, and the insert size was assayed on an Agilent Bioanalyzer 2100 system (Agilent, Santa Clara, CA, United States). The prepared library was sequenced on an Illumina Hiseq 2500. Image analysis and base calling were performed with an Illumina CASAVA pipeline, and finally 125 bp/150 bp paired-end reads were generated.

TABLE 1 | Media formulations at various stages and tissues collected for sequencing.

Stages	Basal medium (pH = 5.8, with 3% sucrose and 7 g/L agar)	Culture time	Materials source
Explants of shoot tips	/	/	Shoot tips
Callus induction	MS + 0.2 mg/L TDZ + 0.6 mg/L 6-BA + 0.15 mg/L 2,4-D + 0.6 mg/L NAA	30 days	Calli
Shoot induction	MS + 1 mg/L 6-BA + 0.1 mg/L NAA	40 days	Leaves
Shoot elongation	MS + 0.1 mg/L NAA + 0.1 mg/L IBA	25 days	Leaves
Rooting	1/2MS + 0.2 mg/L IBA	20 days	Leaves
Outplanting and acclimation	Peat soil:perlite:vermiculite = 3:1:1	30 days	Leaves

We used FastQC (fastqc_v0.11.5) to perform quality control of the raw reads. Then, adapter sequences and low quality reads were removed through Trimmomatic (Trimmomatic-0.36) software using the following parameters (SLIDINGWINDOW: 4:15; LEADING: 3, TRAILING: 3; ILLUMINACLIP: adapter.fa: 2: 30: 10; MINLEN: 36). The remaining reads that passed all the filtering steps were counted as clean reads and all subsequent analyses were based on these data.

Reads Mapping to the Reference Genome

We have performed *de novo* genome sequencing of *F. nilgerrensis* (Qiao et al., 2021). Here, we used Bismark software (version 0.16.3) (Karim et al., 2016) to align the bisulfite-treated reads to our sequenced reference genome. The reference genome was firstly transformed into a bisulfite-converted version (C-to-T and G-to-A converted) and then indexed using bowtie2 (Langmead and Salzberg, 2012). Sequence reads were also transformed into fully bisulfite-converted versions (C-to-T and G-to-A converted) before they were aligned to similarly converted versions of the genome in a directional manner. Sequence reads that produced a unique best alignment from the two alignment processes (original top and bottom strand) were then compared to the normal genomic sequence and the methylation state of all cytosine positions in the read was inferred. The same reads that aligned to the same regions of the genome were regarded as duplicated ones. The sequencing depth and coverage were summarized using deduplicated reads.

Genome-Wide DNA Methylation Distributions Analysis

In order to calculate the methylation level of the sequence, we divided the sequence into multiple bins, with a bin size of 10 kb. The sums of methylated and unmethylated read counts in each window were calculated. Methylation level (ML) for each C site shows the fraction of methylated Cs (mC) and is defined by the following equation: ML = reads (mC)/reads (mC + umC), where umC are the non-methylated Cs.

Calculated ML was further corrected with the bisulfite non-conversion rate according to previous studies (Lister et al., 2013). The calculation was based on the percentage of methylated cytosine in the entire genome, in each chromosome and different regions of the genome, and in three sequence contexts (CG, CHG, and CHH).

Detection of Differentially Methylated Regions and Their Related Genes

Differentially methylated regions were identified using the DSS package (Wu et al., 2015). The DSS method uses spatial correlation (the level of methylation at sites adjacent to cytosine), the sequencing depth of cytosine sites, and the difference between biological repeats to detect and evaluate DMRs. According to the distribution of DMRs through the genome, we defined the genes related to DMRs as genes whose gene body region (from TSS to TES) or promoter region (upstream 2 kb from the TSS) had an overlap with the DMRs.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Enrichment Analyses of Differentially Methylated Region-Related Genes

Gene Ontology (GO) enrichment analysis of genes related to DMRs was implemented by the GOseq R package (Young et al., 2010), in which gene length bias was corrected. GO terms with corrected p-values less than 0.05 were considered significantly enriched by DMR-related genes. The main feature of KEGG (Kyoto Encyclopedia of Genes and Genomes) is to link genes with various biochemical reactions. We used KOBAS software (Mao et al., 2005) to test the statistical enrichment of DMR-related genes in the KEGG pathways. Similarly corrected pathways with p-value < 0.05 were considered to be pathways with significant enrichment of DMR-related genes.

Quantitative Real-Time PCR Validation of Differentially Methylated Region-Related Genes

We randomly selected 20 differentially methylated region-associated genes (DMGs) with significant changes in methylation level at each stage and verified them by qRT-PCR (Quantitative real-time PCR). Total RNA was extracted from six stage samples using a plant RNA Kit (OMEGA bio-tek, Guangzhou, China). Reverse transcription of total RNA was conducted with the PrimeScript RT kit (Takara, Dalian, China) as per the manufacturer's protocol. Each complementary DNA sample was assayed on the QuantStudio 7 Flex real time PCR system software (Thermo Fisher Scientific, United States) with TB green Premix Ex Taq II (Tli RNaseH plus) kit. Gene primers for each gene are listed in **Supplementary Table S1**. The $2^{-\Delta\Delta Ct}$ method

(Livak and Schmittgen, 2001) was employed for normalization of the relative expression of each gene using *FnACTIN* as an internal reference. Each qRT-PCR experiment consisted of three independent biological replicates with two technical replicates for each.

RESULTS

Sequencing Samples and General Evaluation of Whole Genome Bisulfite Sequencing

Different tissues were collected from six stages of the tissue culture process: shoot tips were sampled from explants stage (P1), calli were sampled from the callus induction stage (P2), leaves were collected from shoot induction (P3), shoot elongation (P4), rooting (P5) and outplanting (P6) stages, respectively (**Figure 1** and **Table 1**). Each sample included three biological replicates for Genome Bisulfite Sequencing. The optimized protocol of strawberry micropropagation used in this study is shown in **Table 1**, as developed previously in our lab. A total of 18 samples from six stages of the tissue culture process were collected and sequenced. After three types of cytosine methylation were calculated in the three replicates of each stage, we found that the methylation levels of four samples deviated from other corresponding replicates. Therefore, samples P1-3, P3-1, P5-2, P6-3 were eliminated and a total of 14 samples were used for further analysis. All the Pearson correlation coefficients (R^2) among the replicates were >0.95 in three sequences contexts, indicating high reproducibility between stage-specific replicates (**Supplementary Figure S1**).

A total of 171.44 G raw reads were generated for 14 samples by WGBS. After quality control, 154.38 G clean reads were obtained, with an average of 11 G clean reads per sample; the lowest Q20 and Q30 were 96.72% and 90.25%, respectively (**Supplementary Table S2**). The unique mapping rate of 14 samples ranged from 60.43–75.33%. The average coverage depth of C sites ranged from 7.2×–13.0× (**Supplementary Table S3**). At the same time, the BS conversion rate of the sequencing library was >99.294%, indicating that the DNA methylation information on reads was highly reliable.

DNA Methylation Profiling Varied During Different Stages of Tissue Culture

To comprehensively understand the global DNA methylation dynamics during the tissue culture process of *F. nilgerrensis*, we generated genome-wide methylation profiles of *F. nilgerrensis* for six stages. Global methylation levels showed that CG sites exhibited the highest methylation level in all stages with an average of 49.5%, followed by CHG contexts; CHH contexts were the lowest, with an average of 33.2% and 12.4%, respectively (**Supplementary Table S4**). These differences could be explained by different types of methylation being regulated by different genes (Cokus et al., 2008). Accordingly, the overall distribution of cytosine methylation levels showed that the CG and CHG contexts had greater proportions of higher methylation levels but relatively smaller changes among different stages compared with CHH contexts (**Figure 2A**). This result was

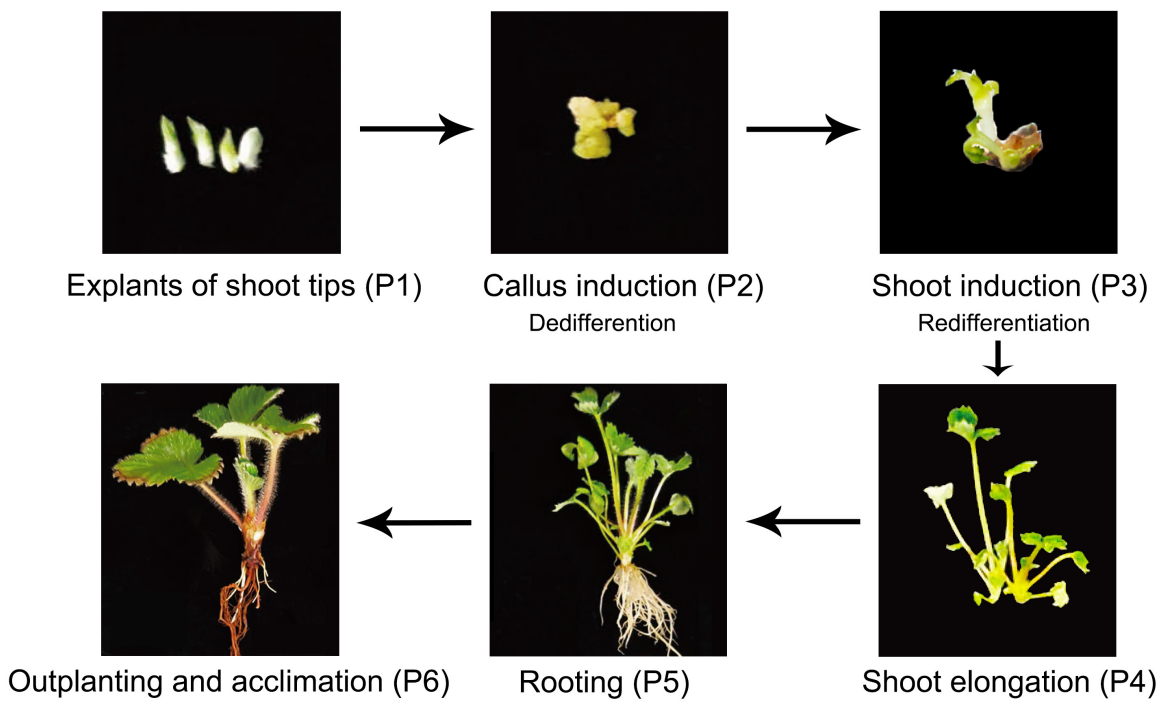

FIGURE 1 | Plant materials used in this study. Plant regeneration of *F. nilgerrensis* from explants of shoot tips to outplanting and acclimation.

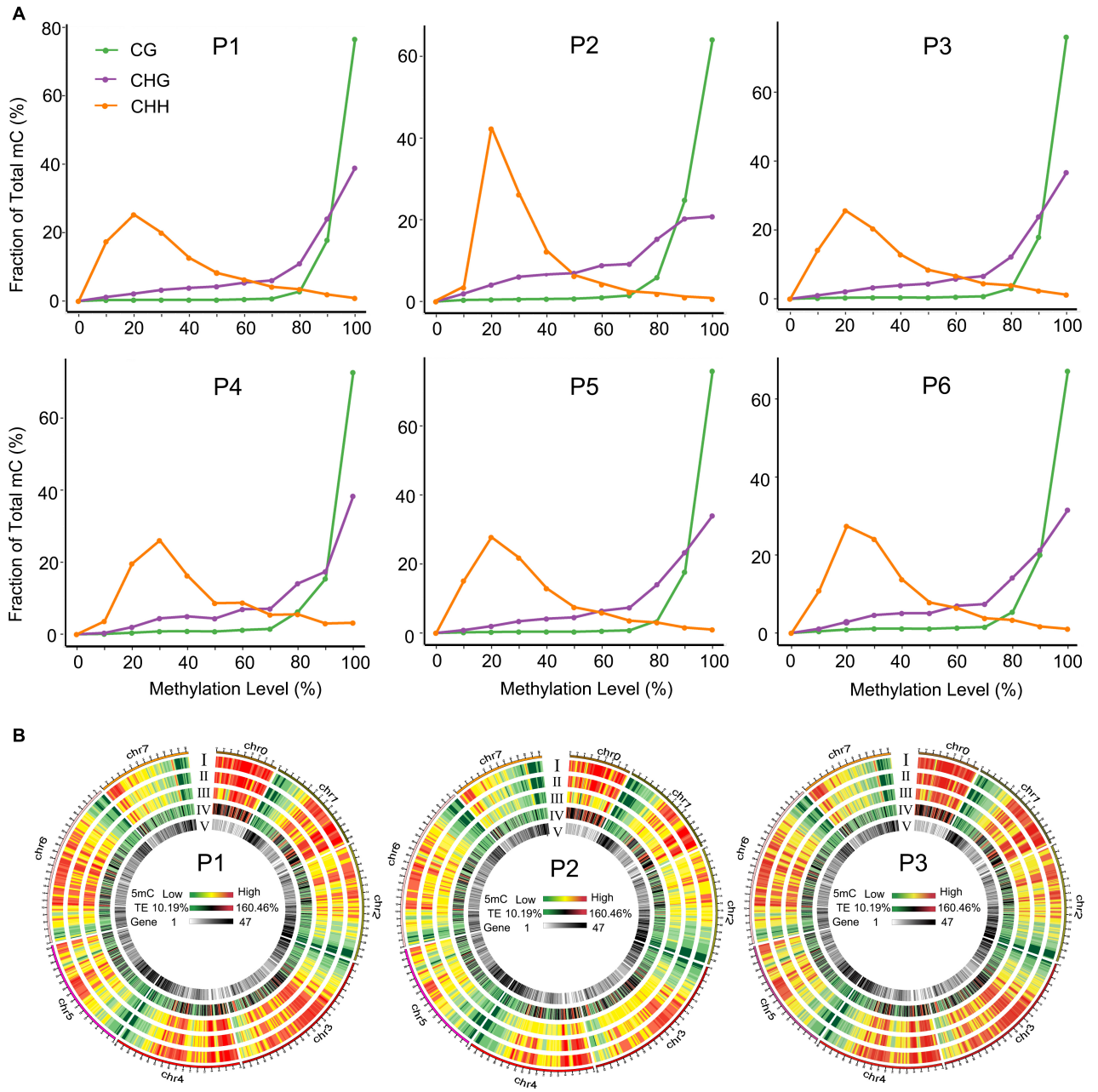

FIGURE 2 | Distribution of methylation levels in different tissue culture stages. **(A)** Distribution of methylation levels in different sequence contexts. **(B)** Circos heat maps of methylation level and gene density distribution along chromosomes of *F. nilgerrensis* in P1 (shoot tips), P2 (callus induction), and P3 (shoot induction). Heatmaps I to V represented the methylation density in the CG, CHG, CHH contexts, TE density, and gene density, respectively.

also supported by the global methylation density distribution map of dedifferentiation and redifferentiation processes (P1–P3), which indicated that the methylation level of CHH contexts obviously changed between different stages and most of these changes occurred in the TE high-density regions (**Figure 2B**). Consistent with the global methylation patterns, the distribution of methylated cytosines along chromosomes was uneven, of which the proportion of CHH methylation changed dramatically (**Supplementary Figure S2**).

The distribution of DNA methylation levels among genetic regions and repeats was also significantly different, e.g., the DNA methylation levels in all three contexts were much higher in repeats, promoters and introns than in the other regions (5′UTR, exons, 3′UTR) in the six stages (**Figure 3A**). Heat map analysis produced a similar pattern, which showed different methylation levels in different gene components (**Figure 3B**). We also found that the methylation levels of P1, P3, and P5 were higher than P2, P4, and P6 stages in each sequence context of

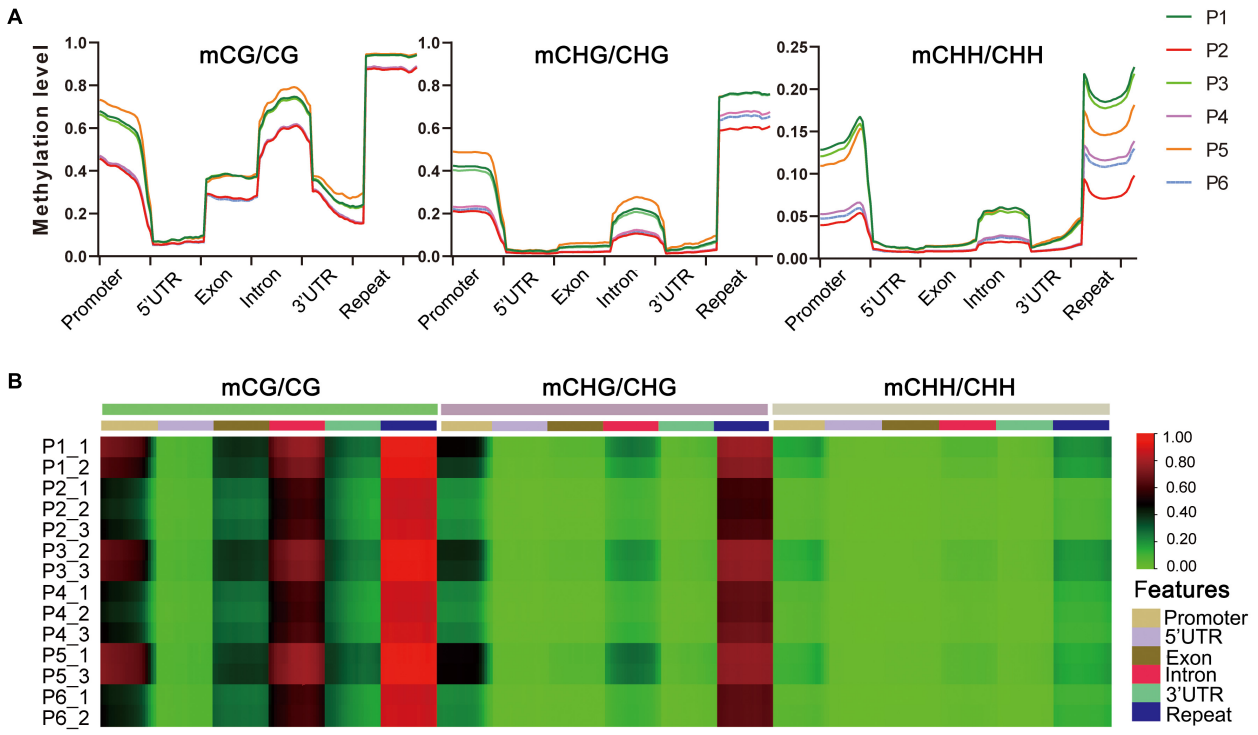

FIGURE 3 | DNA methylation levels in different genomic regions of *F. nilgerrensis*. **(A)** Distribution of DNA methylation among gene component and repeats. **(B)** Heat map of methylation levels of promoter, repeat, coding genes with 5′UTR, exon, intron, and 3′UTR in each sample.

genetic regions and repeats, among which P2 (callus induction) exhibited the lowest methylation levels in all three contexts (**Figure 3A** and **Supplementary Table S4**). This suggested that the dynamically changed cytosine methylation exhibited in the *in vitro* culture of *F. nilgerrensis* may play an important role in regulating gene expression at different stages, leading to varied phenotypic features.

Dynamic Pattern of Differentially Methylated Regions Changes Among the Six Stages

To explore the relationship between DNA methylation and *in vitro* regeneration, we identified DMRs between each adjacent stage. The results showed that many more DMRs were detected in P2 vs. P1 (dedifferentiation) and P3 vs. P2 (redifferentiation) than with other comparisons, suggesting epigenetic regulation plays an important role in reprogramming of gene expression in cell dedifferentiation and redifferentiation. Furthermore, during the tissue culture process, either hypo- or hyper-DMRs alternately dominant (**Figure 4A**).

Consistent with results from the genome-wide methylation profiles map, the CHH sites, where methylation levels altered dramatically, accounted for a larger proportion of DMRs in the first three comparisons (**Figure 4B**), but more DMRs were detected in CG and CHG sites in the last two comparisons, including shoot elongation, rooting, as well as outplanting and acclimation (**Figure 4B**). Further analysis of the CHH methylation distribution on the genetic components indicated that most of them occurred in repeats (most are transposons, TE) and were hypo- and hypermethylated alternatively (**Figure 4C**) among different stages. Different from CHH-DMRs, a large number of CG-DMRs were distributed in exons and promoters, and a large number of CHG-DMRs occupied either repeats or exons at a different developmental switch (**Figure 4C**).

Differentially Methylated Region-Associated Genes and Functional Enrichment Analysis

Differentially methylated region-associated genes were analyzed based on DMRs that overlapped gene functional regions (such as promoters, UTRs, exons, and introns) with at least 1 bp (Chen et al., 2020). We analyzed the distribution of DMGs components and TEs for the different stages. Interestingly, the changing trend of TE, which account for the highest proportion of DMRs and which decreased dramatically after P3 vs. P2, showed an opposite direction to other genetic components (exons, promoters, and UTRs) (**Figure 5A**). That was in accordance with the observation mentioned above that TEs were mostly affected during dedifferentiation and redifferentiation at three contexts.

The DMGs detected in each comparison between adjacent stages ranged from 7317 to 11246, and 2515 DMGs were shared in all five comparisons (**Figure 5B**), suggesting the methylation level of these genes continuously oscillated during different stages of the tissue culture process. To explore the function

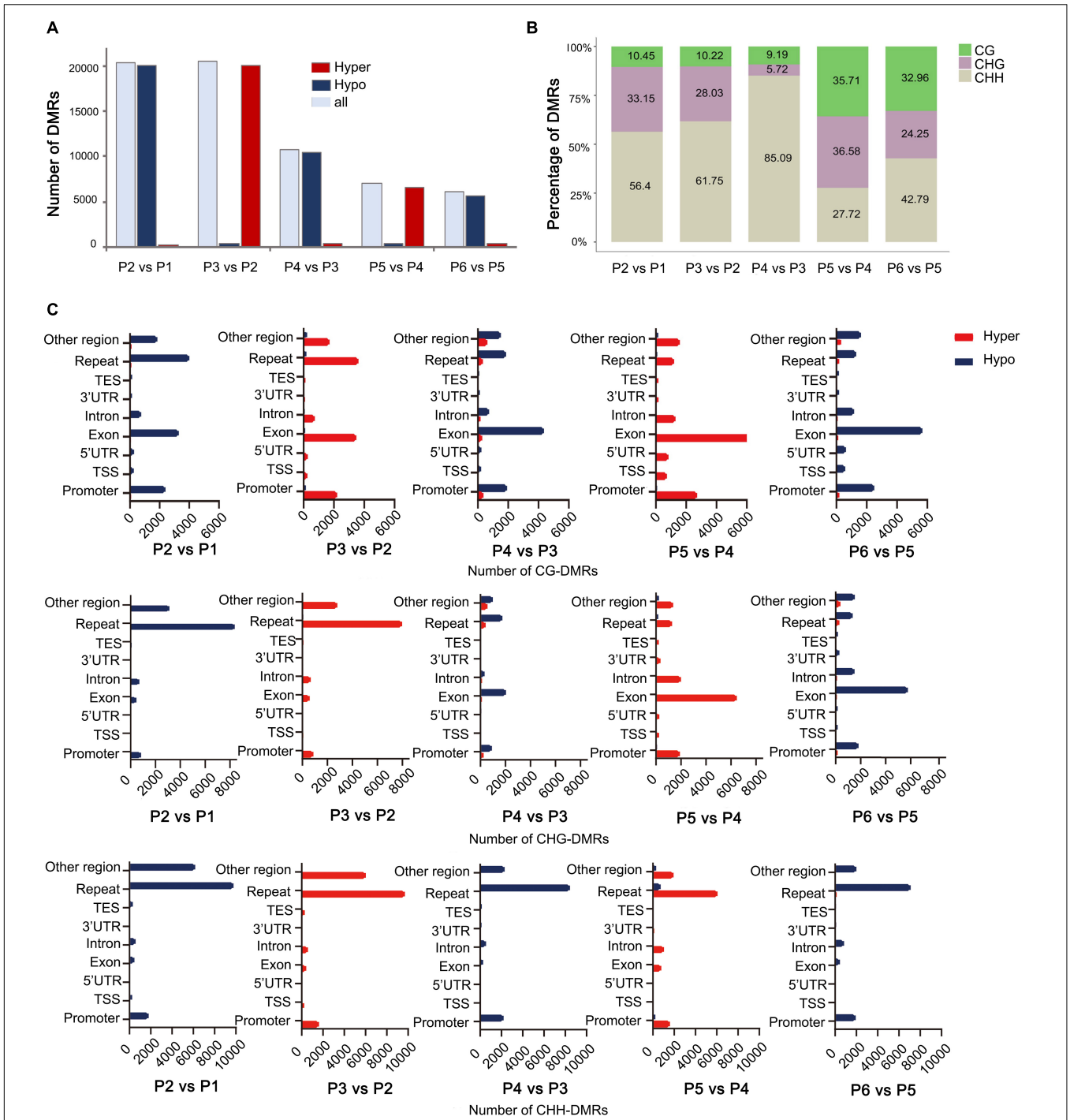

FIGURE 4 | Distribution of DMRs among different stages of tissue culture of *F. nilgerrensis*. **(A)** The number of DMRs in each comparation between adjacent stages. **(B)** Proportion of three contexts (CG, CHG, CHH) in DMRs between each adjacent stage. **(C)** Number of CG/CHG/CHH-DMRs distributed in different genomic regions.

of these DMGs, GO and KEGG enrichment analyses were conducted. The results indicated that the GO terms involved in hormone metabolic processes, plant development and the response to various environmental factors (including bacteria, far red light, hormone, and hypoxia) were enriched throughout the tissue culture process (**Figure 5C** and **Supplementary Table S5**). It was noteworthy that some specific GO terms were also enriched at different stages. DNA methylation-dependent heterochromatin assembly (GO:0006346), embryo development (GO:0009790) and response to wounding (GO:0009611),

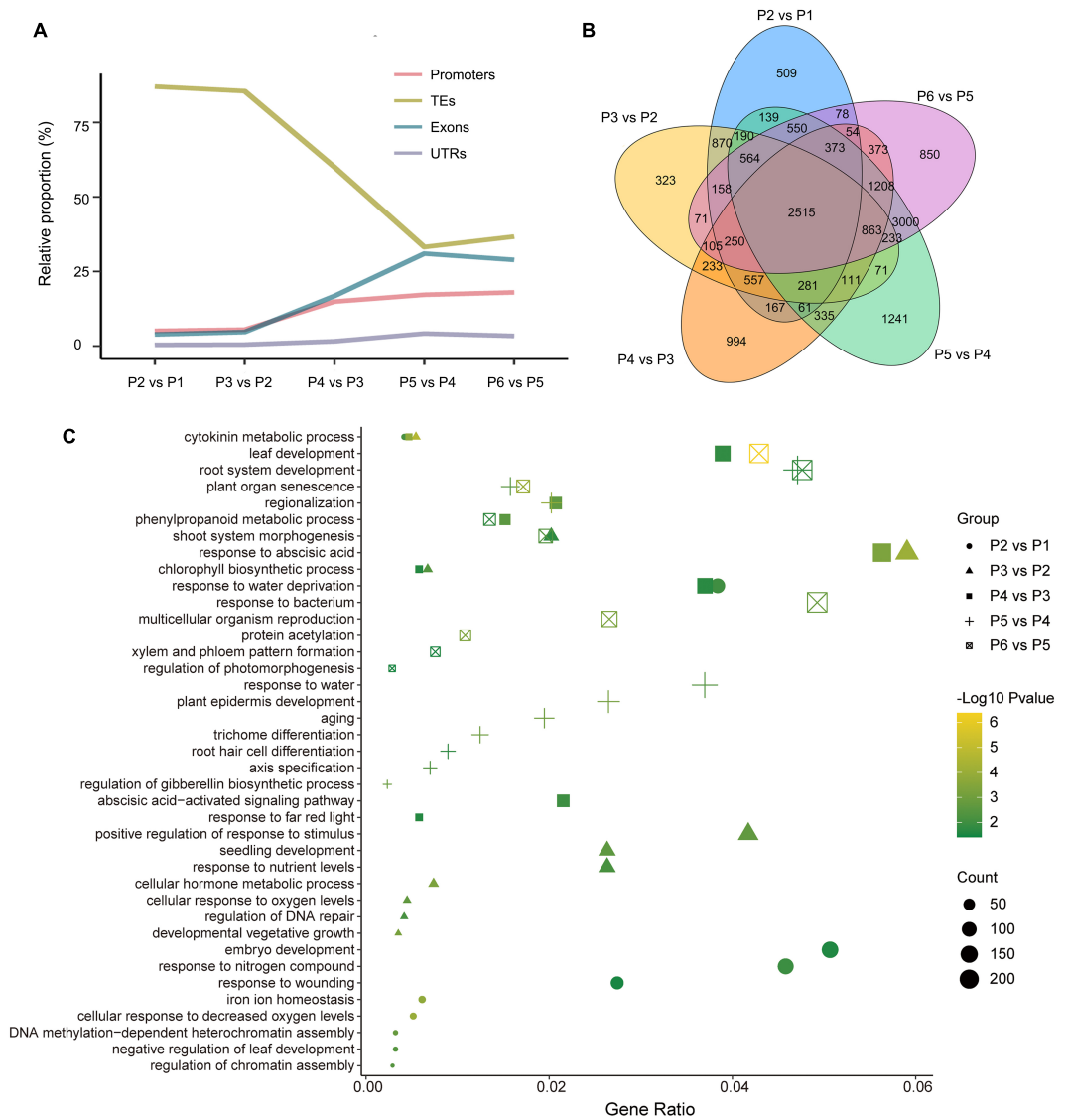

FIGURE 5 | Overview of genome-wide DMR-associated genes (DMGs) during the tissue culture process of *F. nilgerrensis*. **(A)** Dynamic distribution of DMGs in each gene component. **(B)** Venn diagram showing DMGs among the different stage. **(C)** Go enrichment in each comparison between adjacent stages.

shoot system morphogenesis (GO:0010016) were enriched in P1–P3, while regionalization (GO:0003002), phenylpropanoid metabolic processes (GO:0009698), leaf and root development (GO:0048366, GO:0022622), as well as multicellular organism reproduction (GO:0032504) were enriched in the P4–P6 stages (**Figure 5C**).

Correlation Between DNA Methylation and Expression in a Set of Differentially Methylated Region-Associated Genes

Finally, 25 genes, whose methylation levels changed significantly at each stage were listed and expression of 20 genes randomly selected from them were verified by qRT-PCR (**Table 2** and **Figure 6**). Consistent with DNA methylation changes, expression levels of these genes also oscillated during *in vitro* culture, but they only showed relationships with DNA methylation in the first three stages. For example, it is obvious that the most of the promoter hypermethylated genes exhibited reduced expression, including *WIN1*, *WOX13*, *CDK, CKX, RAP, LEC2, bHLH68, ILR1, SAU32, KNAT3*, and *HPSE1*, while three gene bodies (exons and UTRs) hypermethylated genes had an increasing expression trend, including *TCP2, CLV1,* and *CDKF*. No significant expression changes were found in genes after the P4 stage, such as *CDK* and *CKX*, which was consistent with no obvious methylation changes of these genes at these stages. Our findings were roughly consistent with previous reports, that promoter methylation appeared to have a repressive effect on expression, while gene body methylation had a positive effect on expression (Li et al., 2012). Notice that in most of these

genes, no correlation was found between DNA methylation and gene expression after the P3 phase, indicating only a partial role of DNA methylation in regulation of gene expression during latter three stages.

DISCUSSION

Epigenetic mechanisms are highly dynamic events that modulate gene expression of plant developmental processes and respond to environmental abiotic stresses (Orowska, 2021). Analysis of the epigenetic landscape of plant tissue culture processes will help to develop methods for reducing or amplifying the mutagenic and epigenetic effects in tissue culture. We herein investigated the genome-wide methylation patterns and differences at the CG, CHG and CHH sites of six developmental stages of tissue culture in *F. nilgerrensis.*

Global DNA Methylation and Differentially Methylated Regions Detected in Different Stages of Tissue Culture

Among the three contexts, the CG context maintained the highest proportion of total cytosine methylation during the tissue culture process of *F. nilgerrensis*, followed by CHG methylation; the lowest proportion was CHH methylation, mostly distributed in TEs. That is consistent with the previous study in angiosperms that showed CG methylation was the predominant context of DNA methylation, which contributed to more than 50% of total cytosine methylation (Niederhuth et al., 2016). Although CHH methylations accounted for the lowest proportion of total cytosine methylation, they showed the most fluctuating methylation changes among different stages.

Many more DMRs were detected in P2 vs. P1 and in P3 vs. P2, responding to the dedifferentiation and redifferentiation process, respectively, and most of these DMRs were distributed in TEs. That was consistent with previous reports that DNA hypomethylation at the callus stage plays a central role in controlling the activation of the transcription process and the transposition of retrotransposons (Cheng et al., 2006; Fukai et al., 2010). These TE regions were re-methylated in the regenerated plants again for the inhibition of active transposons, which would influence the expression of adjacent genes (Kubis et al., 2003; Zakrzewski et al., 2017). It is noteworthy that the changes of TE proportion and that of other genetic components (exons, promoters and UTRs) in DMRs was opposite after the P3 stage. This indicated that the dedifferentiation and redifferentiation process in the *in vitro* culture systems involved activating or silencing of amounts of transposons, while in other developmental stages, expression of hundreds of genes was epigenetically regulated to control the development of many different cell types.

We found that throughout the tissue culture process of *F. nilgerrensis*, the global pattern of DNA methylation showed dynamic and alternated hypo- and hyper-methylation between each adjacent stage. The dynamics of DNA methylation have already been reported to be an important way to actively reprogram, which plays critical roles in transposon silencing, genome stability and gene expression regulation during cell fate transition in both plants and animals (Feng et al., 2010; Zhang et al., 2010). In plant tissue culture, genome wide hypo- and hypermethylation were predominantly observed during the process of dedifferentiation (callus induction) and redifferentiation (shoots induction), respectively, in a variety of plant species (Neelakandan and Wang, 2012; Ghosh, 2016; Hesami et al., 2020; Lin et al., 2021), that was in accordance with our findings. There is little information concerning alterations in DNA methylation following consecutive stages of tissue culture from explants to outplanting. In *Populus trichocarpa*, the methylomes of explants, calli and regenerated internodes were compared, and the results showed that gene body and transposon 5mC were increased in callus but decreased in regenerated internodes, while promoters 5mC continued to decline among tissues (Vining et al., 2013). Our results showed that methylation levels of all the genetic regions were decreased and increased alternately at the first three stages, roughly corresponding to their three tissues. Furthermore, this trend of alternated hypo- and hypermethylation was continued in the following three stages. It has been reported that different types and concentrations of hormones, together with various stresses and ages of explants would affect the growth and development of culture materials, leading to differences in phenotypes, changing the trend of DNA methylation and induce cell clonal mutations (Law and Jacobsen, 2010; Tiwari et al., 2013). Therefore, the dynamic changes of DNA methylation during *in vitro* culture of *F. nilgerrensis* could be explained by different factors in the microenvironment, such as different types and concentrations of hormones, stages of culture, osmotic stress, light stress, and oxidative stress. In addition, the decreased DMRs in the last three stages indicated that DNA methylation was more stable in the plant tissues with high levels of cell differentiation, suggesting that stage of culture is an important factor affecting DNA methylation levels.

The Genes Affected by DNA Methylation in Tissue Culture

Many genes with differential DNA methylation were detected at each stage of tissue culture in *F. nilgerrensis*. GO and KEGG analysis of these genes showed that genes involved in hormone metabolic processes, plant development and response to various environmental factors were enriched throughout the tissue culture process. That corresponds to the different stresses in the microenvironment, including different types and concentrations of hormones, osmotic stress, light stress, and oxidative stress. For example, the *IQM3* (IQ domain-containing protein) was involved in plant responses to adversity stress, and the expression of the gene is closely related to seed germination (Zhou et al., 2010); *RPS4* was reported as a member of the TIR-NBS-LRR family, which is involved in resistance to bacterial pathogens (Gassmann et al., 2010).

At each stage there were specific enriched GO terms which contained genes playing a crucial role in adaptation.

TABLE 2 | List of important genes of methylation differences at different stages of tissue culture of *F. nilgerrensis*.

Stage	Gene name	Properties	Description	Methy. contexts	Diff. methy.	Methy. region
P2 vs. P1 (hypo)	*WIN1*	AP2/ERF transcription factors	Wound inducing protein	CG	−0.32	Promoter
	WOX13	Wuschel-related homeobox	Somatic embryogenesis	CG	−0.63	Promoter
	AGL	Agamous-like MADS-box protein	Promote the formation of secondary somatic embryos	CG	−0.26	Intron
	CDK	Cyclin-dependent kinase	It can promote the formation of callus when it is rich in auxin	CHG	−0.27	Promoter
	CKX	Cytokinin dehydrogenase	Cell cycle reentry and progression exhibition	CG	−0.4	Promoter
	RAP	Late embryogenesis abundant protein	Late embryonic development protein	CHG	−0.6	Promoter
	TCP2	Transcription factor	Eliminate blade characteristics	CG	−0.32	Exon
	LEC2	Domain-containing transcription factor	Embryo regaining	CG	−0.43	Promoter
	bHLH68	Transcription factor bHLH68	Adjust homeostasis and drought resistance	CHG	−0.3	Promoter
P3 vs. P2 (hyper)	*LEC2*	Domain-containing transcription factor	Embryo regaining	CG	0.45	Promoter
	KLCR1	Kinesin light chain-related	During abiotic stress tune	CHG	0.32	Promoter
	RAP	Late embryogenesis abundant protein	Late embryonic development protein	CHG	0.53	Promoter
	CKX	Cytokinin dehydrogenase	Cell cycle reentry and progression exhibition	CG	0.35	Promoter
	ILR1	IAA-amino acid hydrolase ILR1-like 4	Auxin metabolic process	CHG	0.35	Promoter
P4 vs. P3 (hypo)	*SAU32*	Auxin-responsive protein SAUR32	Auxin reactive protein	CG	−0.72	Promoter
	TIP11	Aquaporin TIP1-1	Participate in drought stress	CG	−0.47	Promoter
	CLV1	Receptor protein kinase CLAVATA1	Maintain the homeostasis of stem cells state	CG	−0.38	Exon
P5 vs. P4 (hyper)	*KNAT3*	Homeobox protein knotted-1-like 3	Heterologous expression promotion somatic embryogenesis	CG	0.38	Promoter
	GAOX	Gibberellin 20 oxidase	Overexpression promotes the production of somatic embryos	CG	0.25	Exon
	CDKF	Cyclin-dependent kinase	Cell cycle regulator	CG	0.31	Exon
	RPS4	RT04_ARATH ribosomal protein S4	Related to resistance to bacterial pathogens	CG	−0.37	Promoter
	HPSE1	Heparanase-like protein 1	Binding growth factor and cytokine regulation binding protein white	CHG	−0.55	Promoter
P6 vs. P5 (hypo)	*SPHK*	Sphingosine kinase	Involved in signal transduction in plant cells guide	CG	−0.62	Exon
	PUB32	U-box domain-containing protein 32	Involved in ubiquitination and protein qualitative interaction	CG	−0.33	Promoter
	IQM3	ARATH IQ domain-containing protein	Young seedlings are closely related to cotyledon expansion	CG	−0.29	Promoter

During callus induction, many genes were hypomethylated in *F. nilgerrensis,* including the key genes *WIN* (Wound-induced protein) and *WOX* (WUSCHEL-related Homeobox) for callus formation. It was reported that *WIN* could induce dedifferentiation and proliferation of cells, and *WOX* could react rapidly to a wound, induce auxin maximization and alter cell fate (Lee and Seo, 2018). During shoot induction, many hypomethylated genes restored methylation, such as *CKX* (Cytokinin dehydrogenase 7) and *ILR1* (IAA-amino acid hydrolase ILR1-like 4), both of which are involved in hormone metabolism. It was reported that cytokinin could promote cell proliferation and shoot induction in the callus (Cortleven et al., 2019), and in *A. thaliana* tissue culture, the absorption of IAA by ILR family mutants is lower than that of the wild type, resulting in shorter hypocotyls and fewer lateral roots (Rampey et al., 2004). Therefore, methylation changes of these genes may reflect crucial roles for regulating the dynamic balance of cytokinin and auxin in *F. nilgerrensis* for shoot induction. For the latter three stages, the key candidate genes with changed methylation were mainly involved in maintaining the steady-state of the stem cell (*CLV1*) (Deyoung et al., 2010; Stahl et al., 2013), regulating the cell cycle (*CDKF*) (Shimotohno et al., 2004; Takatsuka et al., 2009), and participating in ubiquitination and the qualitative interaction of proteins (*PUB32*) (Azevedo et al., 2001; Trujillo, 2018). These genes were closely associated with plant regeneration for stress resistance and development.

It was speculated that DNA methylation affects gene expression by enhancing the binding of certain transcription activators or inhibiting the binding of certain transcription repressors (Zhang et al., 2018). Based on the results of qRT-PCR, we found that they did not exhibit consistent relationships between different genetic regions and different stages. In the first three stages, a negative correlation between DNA methylation and gene expression was found in the promoter, while it

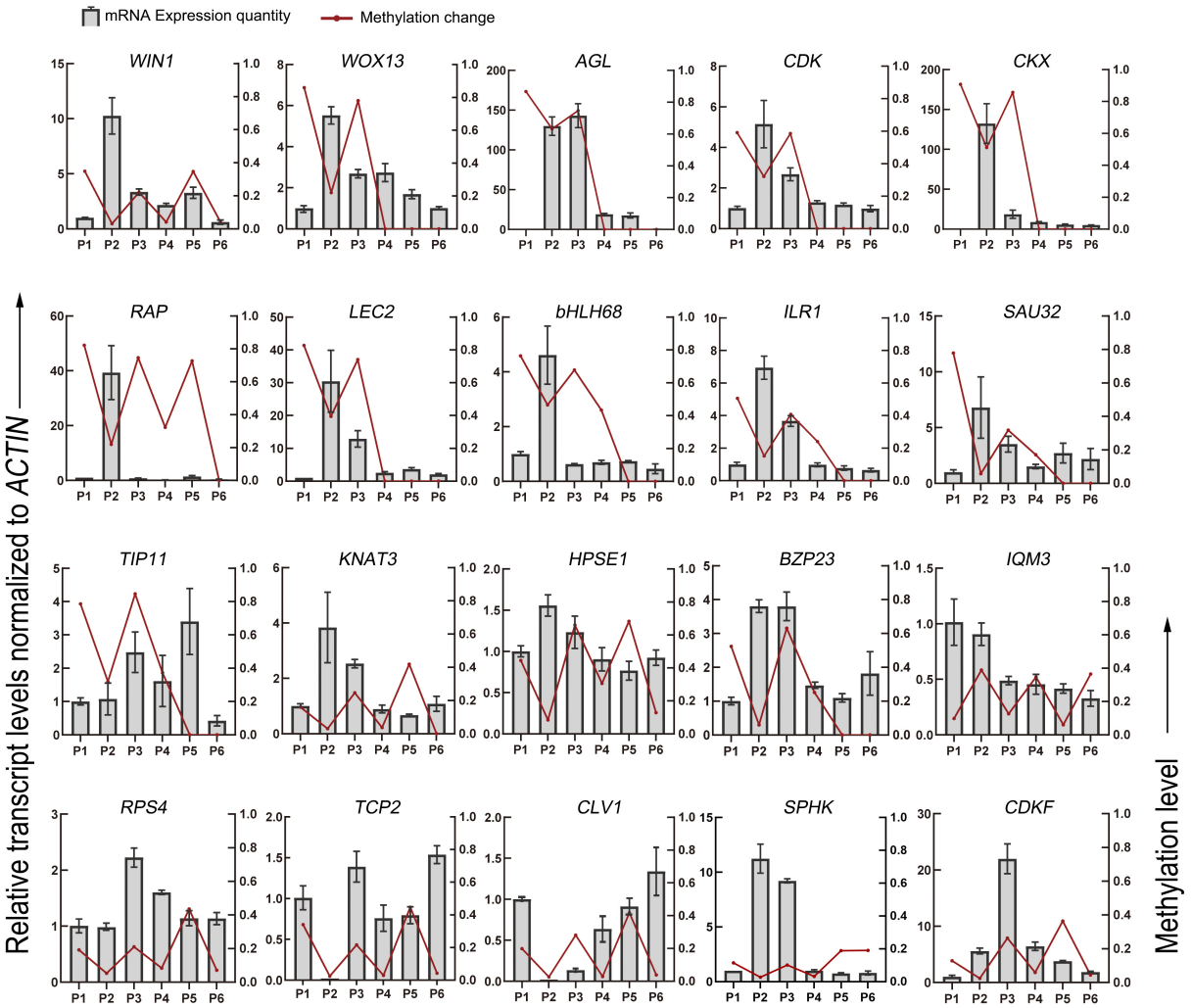

FIGURE 6 | DNA methylation level and gene expression level of DMGs detected in the different stages of tissue culture of *F. nilgerrensis*. The gene expression level was validated by real-time quantitative PCR. The *ACTIN* gene was used as an internal control to standardize the expression of different samples.

seemed that a positive correlation exist in the gene bodies. However, no correlation was detected between methylation and gene expression in the latter three stages. Obviously, several other genetic and epigenetic factors should also be involved in regulating shoot elongation, rooting and outplanting in *F. nilgerrensis*. Our results suggest that the widely accepted belief that hypermethylation leads to repression and hypomethylation leads to activation of genes is an oversimplification, and that this generalization is applicable only in a small fraction of DMGs (Dafni et al., 2018).

In conclusion, we accurately monitored the methylation patterns of consecutive steps of tissue culture by measuring the whole genome DNA methylation levels of *F. nilgerrensis*. We found that the majority of DMRs were located in the TE high-density regions in the dedifferentiation and redifferentiation stages, whereas the proportion of gene-body DMRs gradually increased in the later stages of tissue culture. In addition, we also obtained a series of candidate genes which are closely associated with plant regeneration. This information gives a deeper insight into the relevance of DNA methylation and somatic clonal variation, which can be used to facilitate molecular breeding.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the NCBI Sequence Read Archive (BioProject PRJNA778971), accession numbers SRR16914072, SRR16914071, SRR16914070, SRR16914069, SRR16914068, and SRR16914067.

AUTHOR CONTRIBUTIONS

TZ and QQ conceived and designed the study. QC, YF, XD, LH, and JL performed the experiments and analyzed the data. QC, YF,

PT, MC, TZ, and QQ analyzed the data, wrote the manuscript, and exercised general supervision. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from National Natural Science Foundation of China (32060085 and 31760082 to QQ, 32060237 and 31770408 to TZ) and Applied Basic Research Project of Yunnan (2019FB134 to QQ).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.765383/full#supplementary-material

REFERENCES

Azevedo, C., Santos-Rosa, M. J., and Shirasu, K. (2001). The U-box protein family in plants. *Trends Plant Sci.* 6, 354–358. doi: 10.1016/S1360-1385(01)01960-4

Baránek, M., Křižan, B., Ondrušíková, E., and Pidra, M. (2010). DNA-methylation changes in grapevine somaclones following *in vitro* culture and thermotherapy. *Plant Cell Tissue Organ Cult.* 101, 11–22. doi: 10.1007/s11240-009-9656-1

Chen, Y., Hu, S., Liu, M., Zhao, B., Yang, N., Li, J., et al. (2020). Analysis of Genome DNA Methylation at Inherited Coat Color Dilutions of Rex Rabbits. *Front. Genet.* 11:603528. doi: 10.3389/fgene.2020.603528

Cheng, C., Daigen, M., and Hirochika, H. (2006). Epigenetic regulation of the rice retrotransposon Tos17. *Mol. Genet. Genom.* 276, 378–390. doi: 10.1007/s00438-006-0141-9

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., et al. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452, 215–219. doi: 10.1038/nature06745

Cortleven, A., Leuendorf, J. E., Frank, M., Pezzetta, D., Bolt, S., and Schmülling, T. (2019). Cytokinin action in response to abiotic and biotic stresses in plants. *Plant Cell Environ.* 42, 998–1018. doi: 10.1111/pce.13494

Dafni, A., Anna, E. C., and Francesc, P. (2018). Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenet. Chromatin* 11, 37–54. doi: 10.1186/s13072-018-0205-1

Deyoung, B. J., Bickle, K. L., Schrage, K. J., Muskett, P., Patel, K., and Clark, S. E. (2010). The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in *Arabidopsis*. *Plant J.* 45, 1–16. doi: 10.1111/j.1365-313X.2005.02592.x

Ehrlich, M., and Lacey, M. (2013). DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. *Epigenomics* 5, 553–568. doi: 10.2217/epi.13.43

Feng, C., Wang, J., Harris, A. J., Folta, K. M., and Kang, M. (2020). Tracing the diploid ancestry of the cultivated octoploid strawberry. *Mol. Biol. Evol.* 38, 478–485. doi: 10.1093/molbev/msaa238

Feng, S., Cokus, S. J., Zhang, X., Chen, P. Y., and Jacobsen, S. E. (2010). Conservation and Divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U S A.* 107, 8689–8694. doi: 10.1073/pnas.1002720107

Fukai, E., Umehara, Y., Sato, S., Endo, M., Kouchi, H., Hayashi, M., et al. (2010). Derepression of the plant Chromovirus LORE1 induces germline transposition in regenerated plants. *PLoS Genet.* 6:e1000868. doi: 10.1371/journal.pgen.1000868

Gassmann, W., Hinsch, M. E., and Staskawicz, B. J. (2010). The *Arabidopsis RPS4* bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant J. Cell Mol. Biol.* 20, 265–277. doi: 10.1046/j.1365-313X.1999.00600.x

Ghosh, A. (2016). Detection of DNA methylation pattern in thidiazuron-induced blueberry callus using methylation-sensitive amplification polymorphism. *Biol. Plant.* 61, 511–519. doi: 10.1007/s10535-016-0678-3

Ghosh, A., Igamberdiev, A. U., and Debnath, S. C. (2021). Tissue culture-induced DNA methylation in crop plants: a review. *Mol. Biol. Rep.* 48, 823–841. doi: 10.1007/s11033-020-06062-6

Guo, R., Xue, L., Luo, G., Zhang, T., and Lei, J. (2018). Investigation and taxonomy of wild *Fragaria* resources in Tibet, China. *Genet. Resour. Crop Evol.* 65, 405–415. doi: 10.1007/s10722-017-0541-1

Gupta, S., Pathak, R. U., and Kanungo, M. S. (2006). DNA methylation induced changes in chromatin conformation of the promoter of the vitellogenin II gene of Japanese quail during aging. *Gene* 377, 159–168. doi: 10.1016/j.gene.2006.04.020

Hesami, M., Tohidfar, M., Alizadeh, M., and Daneshvar, M. H. (2020). Effects of sodium nitroprusside on callus browning of *Ficus religiosa*: an important medicinal plant. *J. For. Res* 31, 789–796. doi: 10.1007/s11676-018-0860-x

Horstman, A., Li, M., Heidmann, I., Weemen, M., Chen, B., Angenent, G. C., et al. (2017). The BABY BOOM transcription factor activates the LEC1-ABI3-FUS3-LEC2 network to induce somatic embryogenesis. *Plant Physiol.* 175, 848–857. doi: 10.1104/pp.17.00232

Karim, R., Nuruzzaman, M., Khalid, N., and Harikrishna, J. A. (2016). Importance of DNA and histone methylation in *in vitro* plant propagation for crop improvement: a review. *Ann. Appl. Biol.* 169, 1–16. doi: 10.1111/aab.12280

Kubis, S. E., Castilho, A. M., Vershinin, A. V., and Heslop-Harrison, J. S. (2003). Retroelements, transposons and methylation status in the genome of oil palm (*Elaeis guineensis*) and the relationship to somaclonal variation. *Plant Mol. Biol.* 52, 69–79. doi: 10.1023/a:1023942309092

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

Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719

Lee, K., and Seo, P. J. (2018). Dynamic epigenetic changes during plant regeneration. *Trends Plant Sci.* 23, 235–247. doi: 10.1016/j.tplants.2017.11.009

Li, X., Zhu, J., Hu, F., Ge, S., Ye, M., Xiang, H., et al. (2012). Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* 13:300–315. doi: 10.1186/1471-2164-13-300

Lin, G., He, C., Zheng, J., Koo, D. H., Le, H., Zheng, H., et al. (2021). Chromosome-level genome assembly of a regenerable maize inbred line A188. *Genome Biol.* 22:175. doi: 10.1186/s13059-021-02396-x

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. *Science* 341:21. doi: 10.1126/science.1237905

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, J., Zhang, Y., Diao, X., et al. (2021). Evaluation of Genetic Diversity and Population Structure of *Fragaria nilgerrensis* using EST-SSR Markers. *Gene* 796-797, 145791–145798. doi: 10.1016/j.gene.2021.145791

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: 10.1093/bioinformatics/bti430

Neelakandan, A. K., and Wang, K. (2012). Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. *Plant Cell Rep.* 31, 597–620. doi: 10.1007/s00299-011-1202-z

Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Kim, K. D., Li, Q., et al. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome Biol.* 17:20. doi: 10.1186/s13059-016-1059-0

Noguchi, Y., Mochizuki, T., and Sone, K. (2002). Breeding of a new aromatic strawberry by interspecific hybridization *Fragaria x ananassa* x *F.nilgerrensis*. *J. Jpn. Soc. Hortic. Sci.* 71, 208–213. doi: 10.2503/jjshs.71.208

Orowska, R. (2021). Barley somatic embryogenesis-an attempt to modify variation induced in tissue culture. *J. Biol. Res.* 28, 1–12. doi: 10.1186/s40709-021-00138-5

Qiao, Q., Edger, P. P., Xue, L., Qiong, L., Lu, J., Zhang, Y., et al. (2021). Evolutionary history and pan-genome dynamics of strawberry (*Fragaria* spp.). *Proc. Natl. Acad. Sci.* 118:e2105431118. doi: 10.1073/pnas.2105431118

Rampey, R. A., LeClere, S., Kowalczyk, M., Ljung, K., Sandberg, G., and Bartel, B. (2004). A family of auxin-conjugate hydrolases that contributes to free indole-3-acetic acid levels during *Arabidopsis* germination. *Plant Physiol.* 135, 978–988. doi: 10.1104/pp.104.039677

Shimotohno, A., Umeda-Hara, C., Bisova, K., Uchimiya, H., and Umeda, M. (2004). The plant-specific kinase CDKF;1 is involved in activating phosphorylation of cyclin-dependent kinase-activating kinases in *Arabidopsis*. *Plant Cell* 16, 2954–2966. doi: 10.1105/tpc.104.025601

Stahl, Y., Grabowski, S., Bleckmann, A., Kühnemuth, R., Weidtkamp-Peters, S., Pinto, K., et al. (2013). Moderation of *Arabidopsis* root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes. *Curr. Biol.* 23, 362–371. doi: 10.1016/j.cub.2013.01.045

Takatsuka, H., Ohno, R., and Umeda, M. (2009). The *Arabidopsis* cyclin-dependent kinase-activating kinase CDKF;1 is a major regulator of cell proliferation and cell expansion but is dispensable for CDKA activation. *Plant J.* 59, 475–487. doi: 10.1111/j.1365-313X.2009.03884.x

Tetsu, K., and Akemi, O. (2013). Induction of epigenetic modifications by RNA-directed DNA methylation (RdDM) and their inheritance to the next generation in plants (New gene modification techniques and useful plant production). *Regul. Plant Growth Dev.* 48, 142–147. doi: 10.3389/fmicb.2014.00595

Tiwari, J. K., Chandel, P., Gupta, S., Gopal, J., and Bhardwaj, V. (2013). Analysis of genetic stability of *in vitro* propagated potato microtubers using DNA markers. *Physiol. Mol. Biol. Plants* 19, 587–595. doi: 10.1007/s12298-013-0190-6

Trujillo, M. (2018). News from the PUB: plant U-box type E3 ubiquitin ligases. *J. Exp. Bot.* 69, 371–384. doi: 10.1093/jxb/erx411

Us-Camas, R., Rivera-Solís, G., Duarte-Aké, F., and De-La-Pea, C. (2014). *In vitro* culture: An epigenetic challenge for plants. *Plant Cell Tissue Organ Cult.* 118, 187–201. doi: 10.1007/s11240-014-0482-8

Vining, K., Pomraning, K. R., Wilhelm, L. J., Ma, C., Pellegrini, M., Di, Y., et al. (2013). Methylome reorganization during in vitro dedifferentiation and regeneration of *Populus trichocarpa*. *BMC Plant Biol.* 13:92. doi: 10.1186/1471-2229-13-92

Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., et al. (2015). Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic Acids Res.* 43:e141. doi: 10.1093/nar/gkv715

Xia, L., Han, J. D., Fang, Y. H., Bai, S. N., and Rao, G. Y. (2017). Expression analyses of embryogenesis-associated genes during somatic embryogenesis of *Adiantum capillus-veneris* L. *in vitro*: new insights into the evolution of reproductive organs in land plants. *Front. Plant Sci.* 8:658. doi: 10.3389/fpls.2017.00658

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, 316–328. doi: 10.1186/gb-2010-11-2-r14

Zakrzewski, F., Schmidt, M., Lijsebettens, M. V., and Schmidt, T. (2017). DNA methylation of retrotransposons, DNA transposons and genes in sugar beet (*Beta vulgaris* L.). *Plant J. Cell Mol. Biol.* 90, 1156–1175. doi: 10.1111/tpj.13526

Zhang, H., Lang, Z., and Zhu, J. K. (2018). Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* 19, 489–506. doi: 10.1038/s41580-018-0016-z

Zhang, M., Kimatu, J. N., Xu, K., and Liu, B. (2010). DNA cytosine methylation in plant development. *J. Genet.* 37, 1–12. doi: 10.1016/S1673-8527(09)60020-5

Zhou, Y., Chen, Y., Yamamoto, K. T., Duan, J., and Tian, C. E. (2010). Sequence and expression analysis of the *Arabidopsis* IQM family. *Acta Physiol. Plant.* 32, 191–198. doi: 10.1007/s11738-009-0398-9

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

REVIEW
published: 30 November 2021
doi: 10.3389/fpls.2021.782450

Reprogramming of Histone H3 Lysine Methylation During Plant Sexual Reproduction

Huihui Fang†, Yuke Shao† and Gang Wu**

State Key Laboratory of Subtropical Silviculture, Laboratory of Plant Molecular and Developmental Biology, Collaborative Innovation Center for Efficient and Green Production of Agriculture in Mountainous Areas of Zhejiang Province, College of Horticulture Science, Zhejiang Agriculture and Forestry University, Hangzhou, China

Plants undergo extensive reprogramming of chromatin status during sexual reproduction, a process vital to cell specification and pluri- or totipotency establishment. As a crucial way to regulate chromatin organization and transcriptional activity, histone modification can be reprogrammed during sporogenesis, gametogenesis, and embryogenesis in flowering plants. In this review, we first introduce enzymes required for writing, recognizing, and removing methylation marks on lysine residues in histone H3 tails, and describe their differential expression patterns in reproductive tissues, then we summarize their functions in the reprogramming of H3 lysine methylation and the corresponding chromatin re-organization during sexual reproduction in *Arabidopsis*, and finally we discuss the molecular significance of histone reprogramming in maintaining the pluri- or totipotency of gametes and the zygote, and in establishing novel cell fates throughout the plant life cycle. Despite rapid achievements in understanding the molecular mechanism and function of the reprogramming of chromatin status in plant development, the research in this area still remains a challenge. Technological breakthroughs in cell-specific epigenomic profiling in the future will ultimately provide a solution for this challenge.

Keywords: H3 lysine methylation reprogramming, plant sexual reproduction, histone lysine methyltransferases, histone methylation readers, histone demethylases, *Arabidopsis*

OPEN ACCESS

Edited by:
Paloma Moncaleán, Neiker Tecnalia, Spain

Reviewed by:
Pilar S. Testillano, Margarita Salas Center for Biological Research, Spanish National Research Council (CSIC), Spain
Alexandre Berr, UPR 2357 Institut de Biologie Moléculaire des Plantes (IBMP), France

***Correspondence:**
Huihui Fang
fanghh@zafu.edu.cn
Gang Wu
wugang@zafu.edu.cn

†These authors have contributed equally to this work

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

Received: *24 September 2021*
Accepted: *08 November 2021*
Published: *30 November 2021*

Citation:
Fang H, Shao Y and Wu G (2021) Reprogramming of Histone H3 Lysine Methylation During Plant Sexual Reproduction. Front. Plant Sci. 12:782450. doi: 10.3389/fpls.2021.782450

INTRODUCTION

Histones are the basic packing and organizing proteins in eukaryotic nuclei that package genomic DNA into nucleosomes, the basic repeating structural unit of the higher-order chromatin. The nucleosome comprises 146 base pairs of DNA wrapped in 1.7 superhelical turns around a histone octamer, which contains two copies of each core histone protein, H2A, H2B, H3, and H4. Although the H1 protein itself does not form part of the nucleosome (Luger et al., 1997, 2012), it acts as a linker histone to stabilize inter-nucleosomal DNA. The amino-terminal tails of histone proteins are subject to various types of posttranslational modifications at specific residues, including methylation, acetylation, ubiquitination, phosphorylation, and sumoylation (De Lucia et al., 2008; Liu et al., 2010; Weinhofer et al., 2010; de la Paz et al., 2015; Borg et al., 2020; Ryu and Hochstrasser, 2021). Histone modifications affect nucleosome packaging, and then the chromatin status and gene transcriptional activity depending on the site and degree of specific modification (Bastow et al., 2004; Alvarez-Venegas and Avramova, 2005; Hiragami-Hamada et al., 2016). Dynamic regulation

of histone modification has been shown to be tightly linked to a variety of developmental processes in both plants and animals (Greer and Shi, 2012; Zhao et al., 2019; Cheng et al., 2020).

Plant sexual reproduction involves two major processes: the meiosis and the following fertilization (Wang and Copenhaver, 2018). The reproductive lineage in plants is established late in floral development, which is opposite to the early germline determination during embryogenesis in most animals (Feng et al., 2010). During plant sexual reproduction, a variety of epigenetic memories and chromatin modifications acquired in response to both developmental and environmental cues before the establishment of reproductive lineage need to be reprogrammed to ensure the integrity of genetic information between generations (Borg et al., 2020; Ono and Kinoshita, 2021). Reprogramming of histone methylation during plant sexual reproduction has been shown to be required for resetting the chromatin status toward pluri- or totipotency in gametes and the zygote, thus ensuring to establish new cell fates during plant sexual reproduction (Feng et al., 2010; Gutierrez-Marcos and Dickinson, 2012; Kawashima and Berger, 2014; She and Baroux, 2014; Borg et al., 2020).

In this review, we start by introducing enzymes required for writing, reading, and removing H3 lysine methylation marks, and then we describe their expression patterns in reproductive tissues, including flower bud, inflorescence, anther, stamen, pollen, ovule, embryo, endosperm, and siliques, and finally, we discuss the reprogramming of H3 lysine methylation during plant sexual reproduction and their biological significance on the establishment of new cell fate, with a particular emphasis on the epigenomic resetting of chromatin status toward gamete pluripotency and zygote totipotency.

"WRITERS," "READERS," AND "ERASERS" FOR HISTONE H3 LYSINE METHYLATION

Histone H3 methylation can occur at various sites, but primarily on lysine (Lys, K) residues, and the K4, K9, K27, and K36 residues on H3 tails can be mono-, di-, and/or trimethylated (Liu et al., 2010). H3 lysine methylation is an important and complex epigenetic mark that decorates both transcriptionally silenced and active chromatin status, depending on the specific sites and degrees of methylation (Greer and Shi, 2012; Atlasi and Stunnenberg, 2017; Samo et al., 2021). Typically, H3K4 and H3K36 methylations are linked to the transcriptionally active chromatin status, while H3K9 and H3K27 methylations correlate with heterochromatinization and transcriptional inactivation (Liu et al., 2010; Cheng et al., 2020). The outcomes of H3 lysine methylation are dynamically regulated by "writers," "readers," and "erasers" of histone methylation. The SET (*S*uppressor of variegation, *E*nhancer of *Z*este and *T*rithorax) Domain Group (SDG) proteins are the main histone lysine methyltransferases (HKMTs), serving as "writers" for adding methylation marks to specific lysine residues on H3. Two types of histone demethylases (HDMs), the lysine-specific demethylase 1 (LSD1, or KDM1) homologs and Jumonji C (JmjC) domain-containing proteins (JMJs), act as "erasers" to remove methylation marks (Xiao et al., 2016; Cheng et al., 2020). Additionally, distinct H3 lysine methylation modifications can recruit specific binding effectors, namely "readers," to recognize histone marks and mediate downstream biological events, including chromatin organization and gene transcriptional regulation (Berger, 2007; Zhao et al., 2018). Here, we introduce some most commonly known "writers," "readers," and "erasers" for H3 lysine methylation (**Figure 1**) based on their targeted residues.

"Writers" for Adding Methylation Marks on H3 Lysine Residues

The SDG proteins serve as "writers" for adding H3 lysine methylation marks, and the SET domain is responsible for the catalytic activities of HKMTs. At least 49 putative SET domain-containing proteins have been identified[1] in *Arabidopsis* (Pontvianne et al., 2010; Cheng et al., 2020), and these SDG proteins can be divided into seven classes (Class I to Class VII) based on their domain architecture and/or difference in enzymatic activity, including Class I, the *E*(Z) (*E*nhancer of *Z*este) homologs; Class II, the ASH1 (Absent, Small, or Homeotic discs 1) groups [ASH1 homologs (ASHH) and ASH1-related proteins (ASHR)]; Class III, the Trx (*T*rithorax) groups (TRX homologs and TRX-related proteins); Class IV, ATXR5 (Arabidopsis *T*rithorax-related 5) and ATXR6, which only existed in yeast and plants, are separated from the TRX subfamily and considered to be a newly separated IV subfamily (Zhou et al., 2020); Class V, the SU(VAR)3-9 sub-groups [SU(VAR)3-9 homologs (SUVH) and SU(VAR)3-9 related proteins (SUVR)] (Pontvianne et al., 2010; Cheng et al., 2020). Additionally, a series of proteins containing the split SET domains were incorporated into the SDG proteins and were further classified into VI (SMYD) and VII (SETD) subfamilies (Springer et al., 2003; Ng et al., 2007). The genetic information and classification of SDG proteins are summarized in **Table 1** based on the specific residues on which they act, and their functions in mediating mono-, di-, or tri-methylation modification on H3K4, H3K9, H3K27, and H3K36 sites are discussed subsequently.

H3K4 methylation is mainly enriched in genic regions but depleted in transposons. H3K4me1 accumulates mainly in the transcribed regions, while H3K4me2/3 are enriched in the promoter and 5′ end of the transcribed regions, with H3K4me3 peaking slightly upstream of H3K4me2 (Zhang et al., 2009; Cheng et al., 2020). H3K4me1 and H3K4me2 associate with both active and inactive transcription, whereas H3K4me3 strongly links to transcriptional activation (Cheng et al., 2020). H3K4me1 is highly correlated with CG DNA methylation in the transcribed regions of genes, but H3K4me2/3 and DNA methylation appear to be mutually exclusive. TrxG-SDG proteins function as the main writers for adding H3K4 methylation marks. The TrxG-SDG members ATX1/SDG27 and ATX2/SDG30 are two chromosomal duplications in *Arabidopsis* with divergent functions in catalyzing H3K4 methylation. ATX1/SDG27 has the H3K4 methyltransferase activity and mainly catalyzes H3K4me3 deposition (Alvarez-Venegas et al., 2003), while ATX2/SDG30

[1] www.chromDB.org

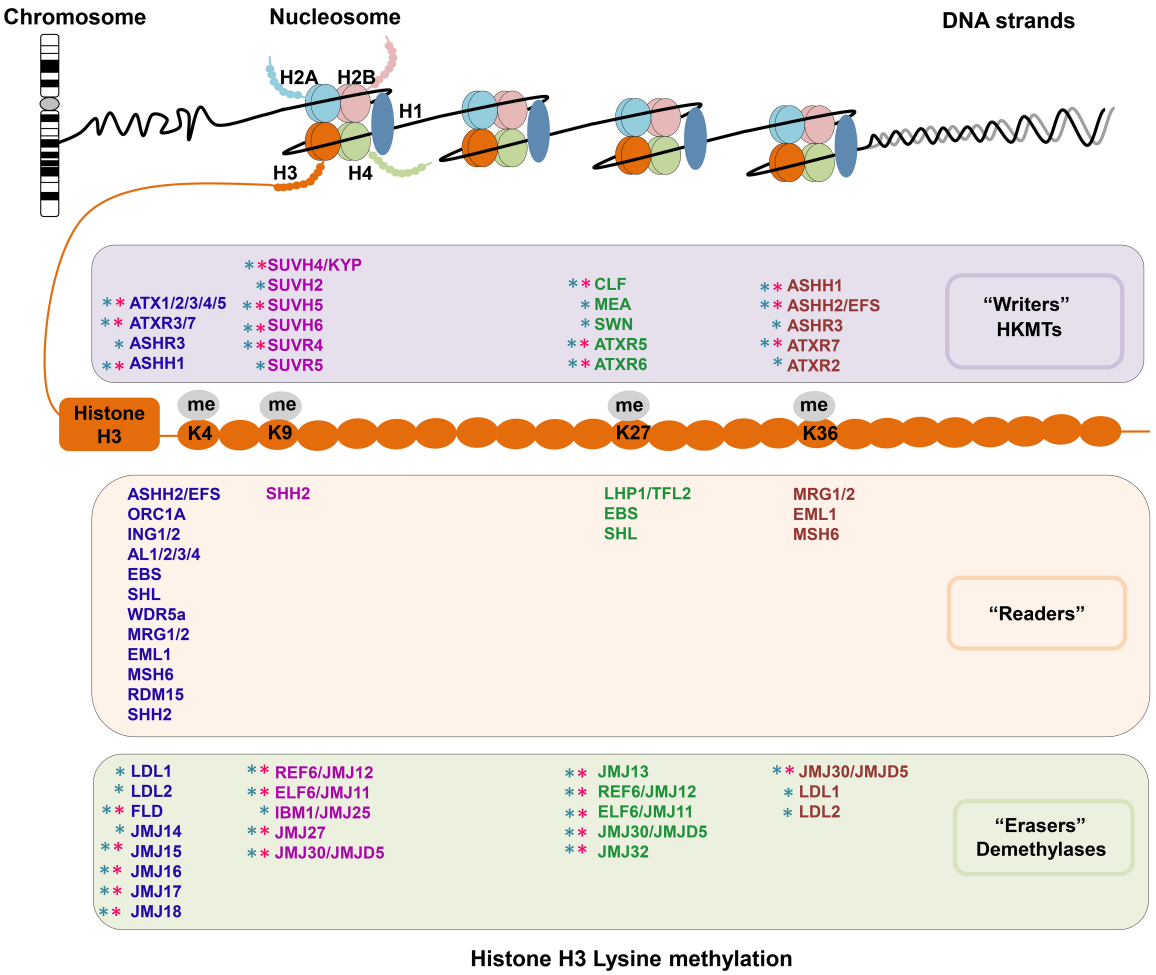

FIGURE 1 | Writers, readers, and erasers of H3 lysine methylation in *Arabidopsis*. The identified histone methyltransferases "writers," binding effectors "readers" and demethylases "erasers" for H3K4, H3K9, H3K27, and H3K36 methylation (me) are summarized in different colored boxes, respectively. The writers are summarized in the light purple box, the readers are summarized in the light orange box, and the erasers are summarized in the light green box. Moreover, H3 methylation associated enzymes acting on different residues are listed in different colors. The writers, readers and erasers for H3K4 methylation are listed in blue words, the enzymes for H3K9 methylation are listed in purple words, the enzymes for H3K27 methylation are listed in green words, and enzymes for H3K36 methylation are listed in brown words. The catalytic specificity for mono-, di-, and trimethylations is not distinguished. Moreover, we added different colored asterisks to discriminate enzymatic activity *in vitro* and deduced from loss-of-function mutants. The pink asterisk represents this enzyme has corresponding enzymatic activity *in vitro*, while the cyan asterisk represents the function of this enzyme was demonstrated by loss-of-function mutant.

is responsible for H3K4me2 (Saleh et al., 2008). This is further confirmed in the *axt1* mutant with reduced H3K4me3 deposition and unchanged H3K4me2 level at the *FLOWERING LOCUS C* (*FLC*) locus. Additionally, ATX1 plays two distinct roles in regulating target genes transcription by facilitating TATA binding protein (TBP) and RNA Polymerase II (Pol II) occupancy at promoters and depositing H3K4me3 within the transcribed region (Ding et al., 2011). ATX3/SDG14, ATX4/SDG16, and ATX5/SDG29 redundantly contribute to both H3K4me2 and H3K4me3 (Chen et al., 2017). ATXR3/SDG2 is the major methyltransferase to catalyze H3K4me3 and plays crucial roles in both sporophyte and gametophyte development (Berr et al., 2010; Guo et al., 2010; Yun et al., 2012; Foroozani et al., 2021). In addition, ATXR3/SDG2 can be co-recruited with a Swd2-like COMPASS axillary subunit onto most transcribed genes to increase H3K4me3 occupancy, which is an atypical and H2B ubiquitination independent pathway (Fiorucci et al., 2019). ATXR7/SDG25 is responsible for all three types of H3K4 methylation (Tamada et al., 2009). In addition, two ASH1-SDG proteins, ASHR3/SDG4 and ASHH1/SDG26 also contribute to H3K4 methylation. ASHR3/SDG4 mediates H3K4me2/3 and H3K36me3 deposition to regulate pollen tube growth and stamen development (Cartagena et al., 2008); ASHH1/SDG26 (Berr et al., 2015) plays roles in both H3K4 and H3K36 methylation. Loss-of-function of ASHH2/EFS/SDG8 reduced H3K4me3 and increased H3K4me2 levels surrounding the CAROTENOID ISOMERASE (CRTISO) translation start site, indicating that ASHH2/EFS/SDG8 regulates H3K4 methylation at least at single or certain genes (Cazzonelli et al., 2009). However, the global profiling of histone methylation demonstrated that the H3K4me3

TABLE 1 | Summary of key writers, readers, and erasers for H3 lysine methylation.

	Writers				Readers			Erasers		
Residue sites	**Formal name**	**ChromDB ID**	**Gene locus**	**Class**	**Protein**	**Gene locus**	**Binding domain**	**Enzymes**	**Gene locus**	**Class**
H3K4	ATX1	SDG27	At2g31650	Class III TrxG-SDG	ORC1A	At4g14700	PHD domain	FLD	At3g10390	LSD1
	ATX2	SDG30	At1g05830		ING1/2	At3g24010/At1g54390		LDL1	At1g62830	
	ATX3	SDG14	At3g61740		AL1/2/3/4	At5g05610/At3g11200/At3g42790/At5g26210		LDL2	At3g13682	
	ATX4	SDG16	At4g27910		EBS	At4g22140	PHD-BAH	JMJ14	At4g20400	KDM5/JARID1-JMJs
	ATX5	SDG29	At5g53430		SHL	At4g39100		JMJ15	At2g34880	
	ATXR3	SDG2	At4g15180		WDR5a	At3g49660	WD40 repeats	JMJ16	At1g08620	
	ATXR7	SDG25	At5g42400		MRG1/2	At4g37280/At1g02740	Chromo domain	JMJ17	At1g63490	
	ASHR3	SDG4	At4g30860	Class II ASH1-SDG	EML1	At3g12140	Tudor domain	JMJ18	At1g30810	
	ASHH1	SDG26	At1g76710		MSH6	At4g02070				
					RDM15	At4g31880				
					ASHH2/EFS/SDG8	At1g77300	CW domain			
					SHH2	At3g18380	Zinc Finger CW domain			
H3K9	SUVH4/KYP	SDG33	At5g13960	Class V Su(var)-SDG	SHH2	At3g18380	Zinc Finger CW domain	IBM/JMJ25	At3g07610	KDM3/JHDM2-JMJs
	SUVH2	SDG3	At2g33290					JMJ27	At4g00990	
	SUVH5	SDG9	At2g35160					JMJ30/JMJD5	At3g20810	JmjC-domain only-JMJs
	SUVH6	SDG23	At2g22740					ELF6/JMJ11	At5g04240	KDM5/JARID1-JMJs
	SUVR4	SDG31	At3g04380					REF6/JMJ12	At3g48430	
	SUVR5	SDG6	At2g23740							
H3K27	ATXR5	SDG15	At5g09790	Class IV	LHP1/TFL2	At5g17690	Chromo domain	ELF6/JMJ11	At5g04240	KDM5/JARID1-JMJs
	ATXR6	SDG34	At5g24330		EBS	At4g22140	PHD-BAH	REF6/JMJ12	At3g48430	
	CLF	SDG1	At2g23380	Class I E(Z)-SDG	SHL	At4g39100		JMJ13	At5g46910	
	SWN	SDG10	At4g02020					JMJ30/JMJD5	At3g20810	JmjC-domain only-JMJs
	MEA	SDG5	At1g02580					JMJ32	At3g45880	
H3K36	ASHH2/EFS	SDG8	At1g77300	Class II ASH1-SDG	MRG1/2	At4g37280/At1g02740	Chromo domain	JMJ30/JMJD5	At3g20810	JmjC-domain only-JMJs
	ASHH1	SDG26	At1g76710		MSH6	At4g02070	Tudor domain	LDL1	At1g62830	LSD1
	ASHR3	SDG4	At4g30860		EML1	At3g12140		LDL2	At3g13682	
	ATXR7	SDG25	At5g42400	Class III TrxG-SDG						
	ATXR2	SDG36	At3g21820	Class VI SMYD						

profiles were comparable between *sdg8* and WT, while H3K36me3 was significantly reduced in *sdg8*, indicating that ASHH2/EFS/SDG8 is mainly responsible for H3K36me3 (McLaughlin et al., 2014; Li et al., 2015).

H3K9me1 and H3K9me2 histone marks are repressive modification marks enriched at chromocenters (Jackson et al., 2004). H3K9me2 is mainly enriched in transposons and repeated sequences, consistent with its primary role in repressing transposon activities and silencing some repeated sequences. Surprisingly, H3K9me3 is mainly associated with euchromatin and transcribed genes, although low levels of this mark are also detected at transposons and repetitive sequences (Mathieu et al., 2005; Charron et al., 2009; Cheng et al., 2020). H3K9 methylation is catalyzed by members of the Su(var)3-9 group proteins. SUVH4/KYP/SDG33 is the first H3K9 methyltransferase to be identified, and it predominantly catalyzes H3K9me2 modification. In *kyp/suvh4* mutant, the deposition of H3K9me2 in heterochromatin is greatly reduced, but H3K9me1 is not affected significantly, implying that SUVH4/KYP/SDG33 has little contribution to H3K9me1 (Liu et al., 2010). SUVH4/KYP/SDG33 is also required for the maintenance of DNA methylation by binding to the methylated cytosines, providing a link between histone modification and DNA methylation (Jackson et al., 2002; Du et al., 2014). SUVH2/SDG3 has no H3K9 methylase activity *in vitro* (Liu et al., 2014), but H3K9me1 level is significantly reduced in the *suvh2* null mutants (Naumann et al., 2005), suggesting an indirect role of SUVH2/SDG3 in regulating H3K9 methylation. SUVH5/SDG9 and SUVH6/SDG23, two homologs of SUVH4/KYP/SDG33, catalyze H3K9me1 and H3K9me2 mainly in transposons and repetitive sequences (Ebbs et al., 2005; Ebbs and Bender, 2006). SUVR4/SDG31 (Thorstensen et al., 2006) and SUVR5/SDG6 (Caro et al., 2012) preferentially mediate the deposition of H3K9me1 and H3K9me2, respectively. The role of other Su(var)3-9 proteins in H3K9 methylation remains elusive.

H3K27 methylation is an important repressive histone modification mark in plants (Johnson et al., 2004; Feng and Jacobsen, 2011). H3K27me1 accumulates significantly in constitutively silenced heterochromatin, which is of great significance in maintaining chromatin structure and transcriptional silencing (Jacob et al., 2009). Moreover, H3K27me1 deposition is independent of DNA methylation, which suggests that H3K27 methylation might have a distinct regulatory mechanism from H3K9 methylation. ATXR5/SDG15 and ATXR6/SDG34, the only two members of Class IV SDG proteins, contribute redundantly to the deposition of H3K27me1 (Jacob et al., 2009). H3K27me2 associates with both euchromatin and heterochromatin (Mathieu et al., 2005; Fuchs et al., 2006), while H3K27me3 is mostly restricted to the transcribed regions of genes to silence a large number of genes in *Arabidopsis* (Zhang et al., 2007a).

H3K27me3 is largely independent of DNA methylation and other epigenetic pathways, but is mainly dependent on Polycomb Repressive Complex 2 (PRC2) (Zhang et al., 2007a; Zheng and Chen, 2011; Xiao and Wagner, 2015). In *Arabidopsis*, members in the PRC2 complex contain three *E*(Z) homologs (CLF/SDG1, MEA/SDG5 and SWN/SDG10), three Su(z)12 homologs (FIS2, EMF2, and VRN2), five p55 homologs (MSI1-5), and only one Esc homolog (FIE) (Kohler et al., 2003). Alternative combinations of these members can form diverse PRC2 complexes, including FIS-PRC2, EMF2-PRC2, and VRN2-PRC2, that play various roles in different developmental processes (Kohler and Villar, 2008; Liu et al., 2010; Golbabapour et al., 2013). PRC1, which acts as a chromatin repressor through mediating the H2A mono-ubiquitination, can communicate with PRC2 (Kahn et al., 2016). VIVIPAROUS (VP1)/ABSCISIC ACID INSENSITIVE 3 (ABI3) LIKE protein (VAL) and Arabidopsis B lymphoma Moloney murine leukemia virus insertion region1 homolog (AtBMI1) mediated H2A ubiquitination initiates the repression of seed maturation genes, and this repression could be further maintained by PRC2-mediated H3K27me3 after the initiation (Yang et al., 2013), suggesting that the PRC1 activity is required for H3K27me3 deposition by PRC2 in certain cases. Another interesting case is that ALFIN1-like (AL) proteins can bind to H3K4me3 marks, and physically interact with PRC1 via its plant homeodomain finger (PHD) domain to form an AL-PHD-PRC1 complex to recruit the PRC2 complex to promote H3K27me3 deposition. This result has important implications for understanding the association between PRC1 and PRC2 complex, as well as the connection between H3K4me3 in gene activation and H3K27me3 in gene repression (Molitor et al., 2014). The bivalent bromo-adjacent homology (BAH)-PHD containing readers capable of recognizing two antagonistic histone marks, including EARLY BOLTING IN SHORT DAY (EBS) and the plant-specific histone reader SHORT LIFE (SHL), regulate floral transition by modulating their binding preference toward either H3K27me3 or H3K4me3 to provide a distinct mechanism of interaction between active and repressive chromatin status (Qian et al., 2018; Yang et al., 2018). Transcription factors (TFs), such as Class I BASIC PENTACYSTEINE (Class I BPC) and C1-2iD ZnF TFs, recruit and interact with PRC2 physically by binding to the Polycomb Response Elements (PREs) (Xiao et al., 2017; Zhou Y. et al., 2018). The telomere-repeat-binding factors (TRBs) recruit CLF/SWN-PRC2 through the telobox-related motifs in *Arabidopsis* (Zhou Y. et al., 2018). A transcriptional repressor, TEMPRANILLO 1 (TEM1), recognizes the 5′-UTR sequence of the *FLOWERING LOCUS T* (*FT*) gene to recruit PRC2 to regulate floral transition (Hu et al., 2021). Recent studies show that two transcriptional repressors, VAL1 and VAL2, are required for PRC2 recruitment for target silencing in *Arabidopsis* (Fouracre et al., 2021; Yuan et al., 2021). TFs with an ethylene-responsive element binding factor-associated amphiphilic repression (EAR) domain can trigger both histone deacetylase complex and PRC2 activities, and different TFs have an additive effect on PRC2 activity (Baile et al., 2021). In the future, it will be of great interest to identify novel proteins required for PRC2 recruitment to a specific locus to regulate gene repression.

H3K36me2 and H3K36me3 are two predominant patterns of H3K36 methylation because H3K36me1 exists simply as a precursor of H3K36me2/3 in plants (Cheng et al., 2020). H3K36 methylation associates with transcriptional activation and transcriptional elongation. H3K36 methylation is specifically mediated by the ASH1-SDG proteins. ASHH2/EFS/SDG8 is

the major H3K36 methyltransferase *in vivo*, and is mainly responsible for H3K36me2/3 (Dong et al., 2008; Xu et al., 2008). It has been reported that ASHH2/EFS/SDG8 can physically interact with the C-terminal domain of the RNA Pol II to facilitate transcription elongation, providing a possible link between Pol II loading and H3K36 methylation deposition (Zhong et al., 2019; Zhang X. et al., 2020). ASHH1/SDG26 has the *in vitro* methyltransferase activity on oligo-nucleosomes, and might repress some gene transcription in an indirect manner (Xu et al., 2008). ASHR3/SDG4 catalyzes H3K36me3 and H3K4me2 to regulate pollen tube growth as their levels were dramatically reduced in the vegetative nuclei in *sdg4* mutant pollen (Cartagena et al., 2008). In addition, although the TrxG-SDG member ATXR7/SDG25 is primarily responsible for H3K4 methylation, it can also catalyze H3K36me2 to activate *FLC* expression to repress flowering (Berr et al., 2009). Moreover, ATXR2/SDG36, a member of the Class VI SMYD subfamily, promotes the accumulation of H3K36me3 during callus formation (Lee et al., 2017).

"Readers" for Recognizing H3 Lysine Methylation Marks

Histone marks can be recognized by specific domains of the effector proteins, referred to as "readers." There are several types of domains that can recognize and bind to H3 lysine methylation marks: the PHD domain, the WD40 repeats (Jiang et al., 2009), and the "Royal Family" domains, which include the Chromo domain, the Tudor domain, the conserved Pro-Trp-Trp-Pro motif (PWWP) malignant brain tumor (MBT) domain, and the plant Agenet module (Zhao et al., 2018). Some readers either contain a single domain, or multiple domains to interact with other factors in macromolecular complexes (Cheng et al., 2020; Xu and Jiang, 2020).

In *Arabidopsis*, a number of proteins have been confirmed to bind to the methylated H3 to generate specific downstream nuclear processes. ORC1, the large subunit of origin-recognition complex (ORC), interacts with the H3K4me3 mark by its PHD finger domain to activate the transcription of target genes (de la Paz and Gutierrez, 2009). The PHD finger containing proteins ING1, ING2, and AL family members, were also shown to be the H3K4me2/3 readers. AL1, AL2, and AL4 have higher affinities to bind to H3K4me3 than H3K4me2, while AL3 has similar binding affinities toward H3K4me2 and H3K4me3 (Zhao et al., 2018). The EBS contains bivalent BAH-PHD reader modules that bind to either H3K27me3 or H3K4me3, and acts as a reader to switch binding between H3K27me3 and H3K4me3, thus timely regulating *FLC* transcription and floral transition (Yang et al., 2018). A plant-specific histone reader SHL can also recognize both H3K27me3 and H3K4me3 via its BAH and PHD domain, and BAH-H3K27me3 and PHD-H3K4me3 interactions are important for SHL-mediated floral repression (Qian et al., 2018).

Arabidopsis WD40-repeat 5a (WDR5a) binds to the K4-methylated H3 tail of FRIGIDA (FRI) specifically to enrich the WDR5a-containing COMPASS-like complex and H3K4 methylation at the *FLC* locus (Jiang et al., 2009). The single chromo domain of LHP1/TFL2 recognizes specifically H3K27me3 marks and represents a potential stabilizing factor of PRC2 activity (Turck et al., 2007; Zhang et al., 2007b). Morf Related Gene (MRG) group proteins, MRG1 and MRG2, bind to H3K4me3 and H3K36me3 peptides through their chromo domains to regulate *FT* transcription and flowering time (Xu et al., 2014; Zy et al., 2014). In *Arabidopsis*, the single Tudor domain protein EMSY-like 1 (EML1) functions as a plant-specific H3K4me2/3 reader, different from the case in humans that only double or tandem Tudor domains can recognize H3K4me2/3 (Zhao et al., 2018), indicating a plant-specific recognition mode. EML1 can also recognize H3K36me3 (Milutinovic et al., 2019) with a much weaker binding affinity for H3K36me3 than for H3K4me3 (Zhao et al., 2018). The Tudor domain protein MSH6, a DNA mismatch repair protein, binds to H3K4me3 with a much weaker affinity than H3K36me3 *in vitro* (Zhao et al., 2018). Recently, it has been reported that RDM15 with a Tudor domain specifically recognizes the H3K4me1 mark, and it functions as an RNA-directed DNA methylation (RdDM) component, thus establishing a link between H3K4me1 and RDM15-mediated RdDM (Niu et al., 2021). The Zinc Finger CW domain of SAWADEE homeodomain homolog 2 (SHH2) has a strong binding affinity for H3K4me3 (Zhao et al., 2018), and the maize SHH2 can specifically recognize H3K9me1 via its SAWADEE domain to establish a functional link between the RdDM pathway and H3K9me1 modification (Wang et al., 2021). The CW domain of ASHH2/EFS/SDG8 exhibits binding preference to H3K4me1, which is different from the mammalian counterpart that has binding preference to H3K4me3 (Liu and Huang, 2018).

Collectively, recognition of a distinct histone mark by a corresponding reader, the histone "mark-reader" pair, indicates a general "*trans*-acting" epigenetic regulatory mechanism in plants (Roudier et al., 2009; Zhao et al., 2018). However, the phenomenon that one reader can simultaneously recognize two or even more histone marks might provide an important mechanism for plants to achieve different biological readouts by modulating the binding affinities of one reader toward multiple marks. Identification of more histone "mark-reader" pairs will be a future challenge.

"Erasers" for Removing Methylation Marks on H3 Lysine Residues

Two types of demethylases, the lysine-specific demethylase 1 (LSD1, or KDM1) homologs and Jumonji C (JmjC) domain-containing proteins (JMJs), contribute to the removal of H3 lysine methylation marks at different sites by interacting with different cofactors (Liu et al., 2010; Xiao et al., 2016). The KDM1/LSD1 demethylases have demethylase activities on di- and mono-methylated lysines, but not on tri-methylated lysines (Klose and Zhang, 2007). The KDM1/LSD1 homologs LDL1, LDL2, and FLD function to remove H3K4me2 marks. Both H3K4me2 and H3K4me3 deposition are elevated at the *FLC* and *FWA* loci in *ldl1ldl2* and *fld* mutants (Jiang et al., 2007; Liu et al., 2007; Shafiq et al., 2014; Berr et al., 2015). Moreover, the increased levels of both H3K4me2/3 and H3K36me3 at *FLC*, *FT*, *MAF2*, *MAF4* and *MAF5* in *ldl1ldl2* mutant suggest

the functions of LDL1 and LDL2 in removing H3K36me3 in addition to H3K4me2/3. A large number of JMJs proteins can act on mono-, di-, and tri-methylated lysines, and they can be divided into five subgroups based on their sequence similarity: the KDM5/JARID1 group, the KDM4/JHDM3 (JmjC domain-containing histone demethylase 3) group, the KDM3/JHDM2 group, the JMJD6 group, and the JmjC domain-only group (Lu et al., 2008; Liu et al., 2010). Among the KDM5/JARID1 proteins, JMJ14/15/16 act on all three types of methylated H3K4 (Lu et al., 2010; Yang et al., 2012; Liu P. et al., 2019; Liu Y. et al., 2019), while JMJ18 can only demethylate H3K4me2/3 (Yang et al., 2012). JMJ27, a member of the KDM3/JHDM2 proteins, demethylates H3K9me1/2 to regulate flowering (Dutta et al., 2017). As a homolog of human KDM3/JHDM2, IBM1/JMJ25 demethylates H3K9me1/2 and prevents the spread of H3K9me2 at loci near TEs and repetitive elements (Saze et al., 2008; Miura et al., 2009). JMJ13, a member in the KDM4/JHDM3 subfamily, acts as an eraser to remove H3K27me3 deposition at the *FT* locus (Zheng et al., 2019). ELF6 and REF6, the other members of this subfamily, erase H3K27me2/3 methylation redundantly during plant development (Yu et al., 2008; Li et al., 2016). In addition, H3K9 methylation status at some loci can also be modulated by *elf6* and *ref6* loss-of-function mutations (Yu et al., 2008), indicating a potential link between the erasure of H3K9 and H3K27 methylation. Two JmjC domain-only proteins, JMJ30 and JMJ32, demethylate H3K27me2/3 jointly at the *FLC* locus to regulate flowering at elevated temperatures (Gan et al., 2014; Crevillen, 2020). JMJ30 can also act as an eraser to remove H3K36me2/3 (Yan et al., 2014) and H3K9me3 marks (Lee et al., 2018), suggesting multiple roles of a single demethylase in various lysine demethylation.

Expression Patterns of "Writers," "Readers," and "Erasers" in Reproductive Tissues

To investigate the effects of these "writers," "readers," and "erasers" on the reprogramming of H3 lysine methylation during sexual reproduction, we performed a thorough analysis of 95 RNA-seq datasets from various reproductive tissues, including flower buds, inflorescence, anther, stamen, pollen, ovule, embryos, endosperm, and siliques (Wolff et al., 2011; Loraine et al., 2013; Willmann et al., 2014; Klepikova et al., 2016; Tedeschi et al., 2017; Pignatta et al., 2018; Rahmati et al., 2018; Zhou M. et al., 2018; Hofmann et al., 2019). We summarized the expression patterns of these genes in reproductive tissues by using available information here, which might contribute to the understanding of their functions during plant reproduction. The datasets were downloaded from the Arabidopsis RNA-Seq Database (ARS) (Zhang H. et al., 2020), and the information of these RNA-seq libraries was summarized in **Supplementary Table 1**. Differential expression patterns of genes encoding "writers," "readers," and "erasers" in different sexual reproductive tissues were shown in the heat map (**Figure 2**), and the relative expression values were presented based on the RNA fragments per kilobase of exon model per million mapped fragments, also known as the FPKM. Some of these genes, including *ATX5*/*SDG29*, *ORC1A*, *JMJ15*, *JMJ32*, and *LDL2* have extremely low or even undetectable expression levels in all these reproductive tissues included, while some readers, including AL family members, EBS, EML1, MRG1, and RDM15, show higher transcript levels in almost all these reproductive tissues. In addition, some genes, including *ASHR3/SDG4*, *ATX3/SDG14*, *ASHH2/EFS/SDG8*, *SUVH6/SDG23*, *ATXR5/SDG15*, *MRG1*, *JMJ13*, *JMJ18,* and *EML1*, have much higher expression levels in pollen than in other tissues. Among them, *ASHR3/SDG4* is the first SDG gene identified to be associated with male sterility in *Arabidopsis*, and it is specifically expressed in open flowers, especially in the pollen, to induce pollen tube elongation. The *ashr3/sdg4* mutant has a larger number of infertile ovules (Cartagena et al., 2008; Zhou et al., 2020). ASHH2/EFS/SDG8 is required for normal anther differentiation, tapetum development and pollen maturation as approximately 90% of the pollen grains were aborted in *ashh2/efs/sdg8* mutant, and the expression of more than 600 genes associated with meiosis, tapetum development, and anther dehiscence was mis-regulated in *ashh2/efs/sdg8* inflorescences (Grini et al., 2009; Zhou et al., 2020). Additionally, ATXR3/SDG2-mediated H3K4me3 plays critical roles in gamete mitotic cell cycle progression and pollen vegetative cell function during male gametogenesis, and it acts indispensably for gametophyte chromatin landscape (Berr et al., 2010; Pinon et al., 2017). Our analysis might imply that the specific expression patterns of these genes in pollen may be of significance for pollen development and function by re-organization of histone modification. It was observed in a previous study that *FLD*-related GUS staining was observed in the anther-filament junction and in the tapetum, but not in the mature pollen grains by using *FLD* promoter-driven GUS transcriptional reporter transgenic lines (Martignago et al., 2019), which is consistent with our analysis that *FLD* is expressed in anther but not in pollen (**Figure 2**). The expression of *SUVH4/KYP/SDG33*, and *ATXR6/SDG34* exhibit extremely low levels in all these reproductive tissues except in developing embryos. Moreover, their expression declines gradually with embryo development, and few transcripts can be detected by RNA-seq in the mature embryos, indicating that some H3 lysine methylation-associated proteins may have more specific and crucial functions at a certain developmental stage. *SWN*, *AL1*, *AL2,* and *AL3* have higher expression levels in endosperm, which is a key evolutionary innovation of flowering plants, and has been identified as the site of genomic imprinting (Gehring and Satyaki, 2017). Correspondingly, the higher expression of SWN might function importantly in maintaining genomic imprinting. MEA/SDG5 has been well characterized, and it is a self-controlled imprinting gene to produce a cascade of parent-specific gene expression (Macdonald, 2012). Although transcriptomic data from central cells were not included in our RNA-seq dataset, studies have shown that MEA is specifically expressed in the late stage of the central cell (Luo et al., 2000), which is the second female gamete to initiate the endosperm lineage after fertilization. MEA loss-of-function leads to a large number of central cells proliferated excessively under unfertilized conditions, resulting in *mea* seeds with only endosperm but without embryos (Schmidt et al., 2013). Endosperm transfer cells

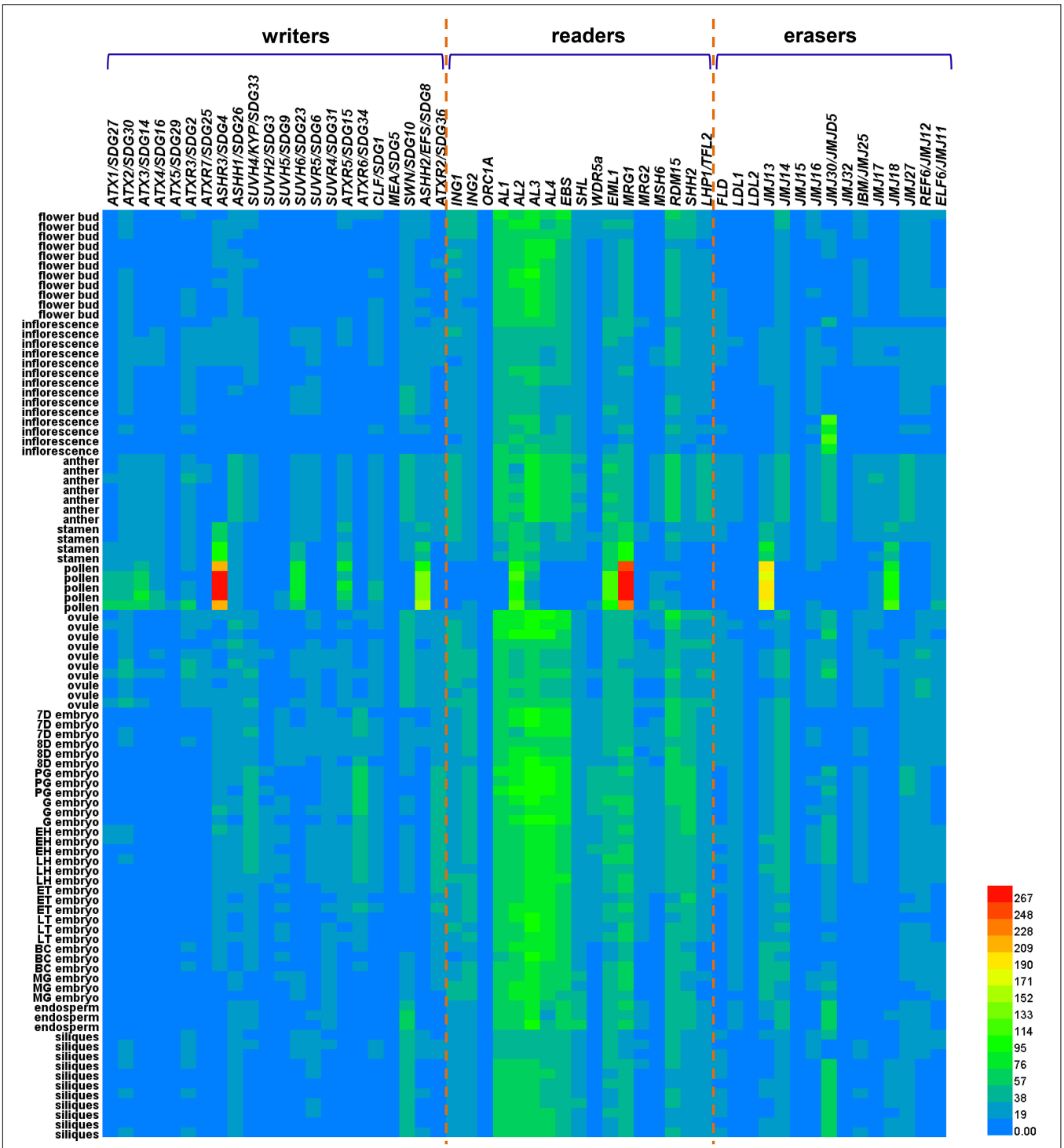

FIGURE 2 | Heat map of differential expression patterns of H3 lysine methylation associated genes across various reproductive tissues in *Arabidopsis.* A total of 95 RNA-seq datasets from various tissues at different stages of plant sexual reproduction, including flower buds, inflorescence, anther, stamen, pollen, ovule, embryos at different developmental stages, endosperm, and siliques were collected and thoroughly analyzed and summarize here. The heat map shows the mRNA expression profile of genes, associated with writing, reading, and erasing the H3 lysine methylation marks during plant sexual reproduction. The symbols of genes are shown at the top of the heat map, and the tissue information are listed at the left side of the heatmap. The bar at the bottom right of the heat map represents relative expression values, which was presented based on the RNA fragments per kilobase of exon model per million mapped fragments, also known as the FPKM. 7D embryo, 7-day-old embryo; 8D embryo, 8-day-old embryo; PG embryo, Preglobular embryo; G embryo, Globular embryo; EH embryo, Early heart embryo; LH embryo, Late heart embryo; ET embryo, Early torpedo embryo; LT embryo, Late torpedo embryo; BC embryo, Bent cotyledon embryo; MG embryo, Mature green embryo. The red and blue denote highly and weakly expressed genes, respectively.

(ETC) are one of the four main types of cells in the endosperm, and *AL1*, another gene highly expressed in the endosperm, has been shown to be an ETC-specific histone reader in rice (Kuwano et al., 2011; Lopato et al., 2014). Collectively, the tissue specific expression patterns of these genes imply their functional importance in these specific tissues. Moreover, it also indicates, from another perspective, that different H3 modification states and reprogramming in different tissues during sexual reproduction might have specific regulatory patterns by distinct enzymes. How these differential expression patterns of certain genes in different reproductive tissues link to their function in reprogramming of H3 lysine methylation during plant sexual reproduction remains an interesting research task.

REPROGRAMMING OF H3 LYSINE METHYLATION DURING PLANT SEXUAL REPRODUCTION

Haploid gametophyte generation in floral organs and the subsequent fertilization have great significance in the alternation of higher plant life cycle and the transgenerational transmission of genetic information (Dahia et al., 2020). Gametophyte generation, including sporogenesis and gametogenesis, involves a series of cell division and differentiation, and the consequent zygote resulted from fertilization has the totipotency to develop into a future seedling. Consequently, plants undergo global chromatin re-organization during sexual reproduction to develop into highly distinct cell types and establish cell pluri- or totipotency, in which the reprogramming of histone modifications plays a vital role (Kawashima and Berger, 2014; She and Baroux, 2014).

Plant sexual reproduction consists of three different phases: sporogenesis, gametogenesis, embryo- and endosperm-genesis. Plant reproduction initiates with sporogenesis, and it is characterized by the generation of meiotic-competent spore mother cells (SMCs), namely SMC differentiation (Kawashima and Berger, 2014; She and Baroux, 2014). The male SMCs, also known as pollen mother cells (PMCs) (*2n*), are differentiated in the sporangium and formed in the anther locule, then undergo meiosis to give rise to four haploid microspores (*1n*). After an asymmetric and atypical mitosis division, each microspore produces one vegetative cell with a larger nucleus (*1n*) and one generative cell with a smaller nucleus (*1n*). Subsequently, the generative cell (*1n*) undergoes one additional mitotic division to generate two sperm cells (*1n*). Mature pollen grain usually contains two sperm cells (*1n*) and one much larger vegetative cell (*1n*) (Twell, 2011). The female gametogenesis begins with the differentiation of the female SMCs, also called megaspore mother cells (MMCs), which occurs within the ovule primordia in the gynoecium. The MMCs (*2n*) undergo meiosis to generate four haploid spores (*1n*), while only one spore survives to form the functional megaspore cell (FMC) (*1n*). The FMC (*1n*) then undergoes three rounds of mitosis to develop into the eight-nucleated mature female gametophyte (embryo sac) consisting of one egg cell (*1n*), one central cell (*2n*), three antipodals (*1n*), two synergids (*1n*) (Drews and Koltunow, 2011; Baroux and Autran, 2015). All cells in the embryo sac are haploid except for the central cell that has a di-haploid maternal genome by inheriting two polar nuclei. Double fertilization is a unique fertilization feature for the angiosperms, and it is a process that two female gametes in the embryo sac, the egg cell (1*n*) and the central cell (2*n*), receive two sperm cells (1*n*) in a one-to-one manner to yield the diploid embryo and the triploid endosperm, respectively. These two fertilization products have distinct developmental fates (Bleckmann et al., 2014). The pre-embryo engages in a series of cell divisions to establish a mature embryo with the potential to develop into a future seedling, while the primary endosperm cell engages in a syncytial phase of proliferation to form an extra-embryonic nurturing tissue.

Thereinafter, we will discuss the reprogramming of H3 lysine methylation and its molecular significance during sporogenesis, gametogenesis, embryo- and endosperm-genesis in flowering plants, mainly in *Arabidopsis*. The main process of sexual reproduction in plants and the dynamic regulation of H3 lysine methylation during these reproductive events are briefly summarized in **Figure 3**.

Reprogramming of H3 Lysine Methylation During the Differentiation of Spore Mother Cells

The earliest event that occurs in sporogenesis is spore mother cells (SMCs) differentiation, which is usually characterized by histone modification-mediated chromatin reprogramming to contribute to the somatic-to-reproductive transition and the meiotic-competence cell fate establishment (Wang and Kohler, 2017; Gehring, 2019).

The differentiation of MMCs, the female SMCs, is marked by nuclear enlargement and chromatin decondensation, and this event coincides with a 60% reduction in heterochromatin content, a decreased number of chromocenters, and the depletion of canonical linker histone (She et al., 2013). Reprogramming of H3 lysine methylation contributes to the formation of the MMC-specific chromatin status, which helps to establish a transcriptionally permissive chromatin status (She et al., 2013). The MMC exhibits a 2.7-fold enrichment of the permissive-associated H3K4me3 mark and a 50% reduction of the repressive-related H3K27me3 mark (Berger and Twell, 2011; She et al., 2013), and some other repressive marks including H3K27me1 are also decreased in the MMC (Berger and Twell, 2011; She et al., 2013; She and Baroux, 2014). The reduction in H3K27me3 (relative to the increase of DNA content) at this pre-meiotic S-phase might be attributable to the non-methylated H3K27 residues in the newly generated nucleosomes by DNA replication. The increase of H3K4me3 mark during the S-phase and prophase I might activate some chromatin-modifying enzymes (She and Baroux, 2014). Strangely, the H3K4me2 level is reduced by 30% and the global H3K9me2 level has a 1.6-fold increase in the MMC (She et al., 2013), this seems inconsistent with the establishment of a permissive chromatin status in MMC as H3K4me2 is an active transcriptional mark while H3K9me2 is a repressive one.

As potentially mobile sequences within the genome, transposable elements (TEs) make it difficult for plants to

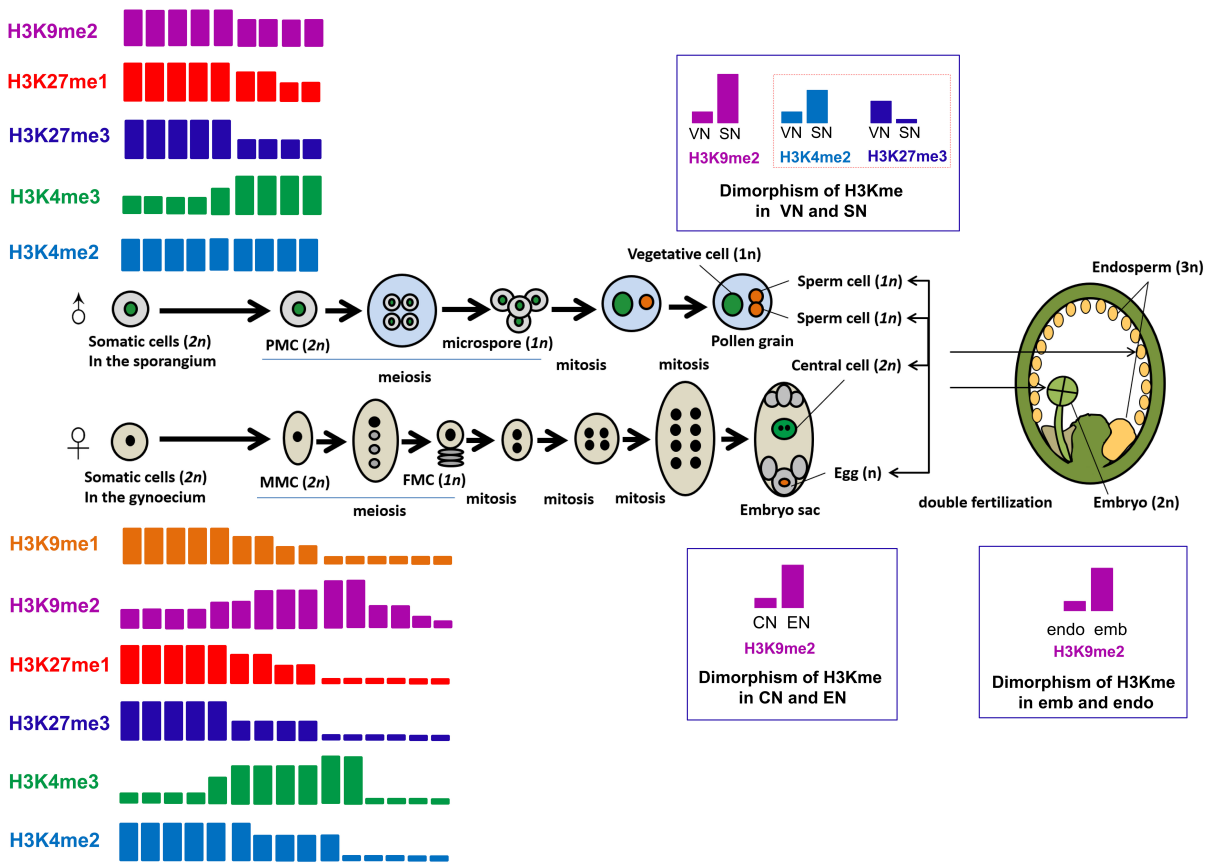

FIGURE 3 | Dynamics of H3 lysine methylation during plant sexual reproduction. The process of plant sexual reproduction, including sporogenesis, gametogenesis, embryo- and endosperm-genesis, as well as the dynamics of H3 lysine methylation during plant sexual reproduction, are summarized in this figure. Differently colored rectangles represent different H3 methylation marks, including the orange rectangle for H3K9me1, the purple rectangle for H3K9me2, the red rectangle for H3K27me1, the blue rectangle for H3K27me3, the green rectangle for H3K4me3, and the light blue for H3K4me2. The height of the rectangles represents the relative enrichment of a certain methylation mark. In the flower, pollen mother cells (PMCs) (*2n*), differentiated from somatic cells in the sporangium, undergo meiosis to form four microspores (*1n*). Each microspore undergoes an asymmetrical division to generate one vegetative cell and one generative cell, which then divides to form two sperm cells (*1n*). The pair of sperm cells and the vegetative cell are linked to form the mature pollen grain. The sperm nucleus (SN) and vegetative nucleus (VN) have dimorphic pattern in their chromatin condensation, and the chromatin of SN is more compacted, accompanied by much more deposition of H3K9me2, compared with that in VN. Paradoxically, the transcriptionally permissive mark H3K4me2 is enriched and the repressive mark H3K27me3 is globally depleted in SN. The megaspore mother cells (MMCs) (*2n*) are generated from somatic cells in the gynoecium. MMCs undergo meiosis to generate four haploid spores (*1n*). The surviving megaspore, functional megaspore cell (FMC) (*1n*), undergoes three rounds of nuclear divisions to generate an eight-nuclei containing syncytial female gametophyte (embryo sac). After cytokinesis, the mature female gametophyte consists of the egg cell, the central cell, and accessory cells (antipodals and synergids). The egg nucleus (EN) exhibits more pronounced condensation, with significantly higher H3K9me2 deposition, compared to the central nucleus (CN). The egg cell and the central cell are each fertilized by one sperm cell to produce the diploid embryo (2n) and the triploid endosperm (3n), respectively. The dimorphic chromatin status present between the embryo (emb) and the endosperm (endo), which might be inherited from the female gametes after fertilization, with the embryo (emb) having highly condensed chromatin compared to the endosperm (endo), characterized by obviously lower H3K9me2 accumulation in emb than in endo.

transmit the genetic information accurately to the next generation. TEs are typically silenced by various repressive machineries including epigenetic modification. However, chromatin decondensation, heterochromatin reduction, and epigenetic reprogramming during plant sporogenesis provide a favorable probability for TEs to escape from silencing; therefore, it is of significance for plants to employ a series of strategies to restrict TEs movement, particularly in the germline (Bao and Yan, 2012). H3K9me2 is an important heterochromatic mark and plays a vital role in silencing the activities of TEs. During the transition of somatic cells to MMCs, H3K9me2 remains highly accumulated in chromocenters (She et al., 2013), suggesting a reinforcement of TEs silencing even though the heterochromatins are not maintained. Moreover, the increase in H3K9me2 levels might be highly specific as H3K27me1, another mark typically enriched at chromocenters, is reduced dramatically in the MMC heterochromatin. H3K9me2 seems to be accumulated by consuming H3K9me1 as H3K9me1 levels decrease in the MMC chromatin (She et al., 2013).

The development of male reproductive lineage begins with the differentiation of PMCs in the early anther locule. In *Arabidopsis*, one sub-epidermal somatic cell in the sporangium enlarges to form an archesporial cell, which then divides to generate a primary sporogenous cell, and subsequently, the

sporogenous cell undergoes mitosis to give rise to PMCs. During PMC differentiation, nuclear morphology undergoes similar changes to that during MMC generation: the PMC exhibits a fivefold increase in nuclear volume size, accompanied by a decrease in heterochromatin content and the average number of distinct chromocenters. This indicates that a distinct nuclear organization related to the transcriptionally permissive chromatin landscape is developed in the PMC (She and Baroux, 2015). Specific histone modification reprogramming occurs to establish this distinct chromatin status in PMCs. The levels of two repressive marks H3K27me1 and H3K27me3 are decreased, while that of the permissive mark H3K4me3 has a 1.8-fold increase in PMCs compared to that in somatic cells. During the PMC differentiation, the level of H3K4me2 is constant compared to that in the surrounding somatic cells (She and Baroux, 2015), different from the decreased H3K4me2 level in MMCs. In PMCs, substantial normally silenced TEs become transcriptionally activated (Chen et al., 2010; Yang et al., 2012), indicating that the decondensation at heterochromatin loci can release some TEs silencing. H3K9me2 levels might be reduced before meiosis in PMCs as it acts to repress TE expression, which is different from the changes of H3K9me2 levels in MMCs. More detailed investigations remain necessary to reveal the dynamic events of histone lysine methylation underlying PMCs differentiation.

Collectively, both the female and male SMCs, MMCs, and PMCs, are competent in differentiating into several distinct cell types. Modulation of H3K27me1, H3K27me3, and H3K4me3 levels in mutants with altered gametophytic competence demonstrates the importance of H3 lysine methylation in pluripotency establishment during SMCs generation (Berr et al., 2010; Olmedo-Monfil et al., 2010). Moreover, the reprogramming of H3 lysine methylation might be a prerequisite for the subsequent meiosis as dynamic regulation of H3 lysine methylation at certain sites is critical for meiotic events, including homologous chromosome pairing, synapsis, and recombination initiation (She et al., 2013). Alternatively, the H3 lysine modification dynamic events might also contribute to activating the meiotic genes and repressing the mitotic pathway.

Reprogramming of H3 Lysine Methylation During Gametogenesis

The female gametogenesis begins with meiosis initiation, then the formation of the eight-nuclei syncytium, and finally, the haploid egg cell and diploid central cell containing mature embryo sac are set definitively by cellularization. Meiotic execution requires additional dynamic histone modifications, particularly during prophase I with further enrichment of H3K4me3 marks occurring along the entire chromosomes (Baroux and Autran, 2015). Moreover, H3K27me1 level in the FMC chromatin is dramatically reduced, while H3K27me3 decreases to an undetectable level (She et al., 2013; Baroux and Autran, 2015). H3K9me2 level shows a more pronounced increase at prophase I during MMC differentiation, while decreases significantly in the FMC chromatin, implying that FMCs go through another wave of H3K9me2 organization (She et al., 2013). During three mitotic cycles of FMCs, the H3K9me2 mark is re-established to a higher level (Pillot et al., 2010).

Currently, histone modification reprogramming events underlying male gametogenesis are barely known, yet a dimorphic chromatin status is established between the sperm cells and the vegetative cell. The sperm cells have highly condensed chromatin, while the vegetative cell has highly decondensed chromatin (Schoft et al., 2009; She and Baroux, 2015).

Dimorphic H3 Lysine Methylation States in Gametes and Their Companion Cells

As companion cells of gametes, the vegetative and central cells have largely decondensed chromatin, which functions importantly in maintaining the integrity of the germline genome and assisting the subsequent fertilization (Ibarra et al., 2012; Baroux and Autran, 2015). Moreover, the companion cells and their corresponding gametes are also marked with a stark dimorphism of the chromatin and transcriptional status, which involves not only unequal DNA methylation (reviewed in Han et al., 2019) but also a dimorphic H3 lysine methylation state. The companion cells have large decondensed chromatin, accompanied by an increase of transcriptionally active histone marks and a reduction of transcriptionally inactive histone marks (Pillot et al., 2010). Correspondingly, the repressive histone marks, including H3K9me2 deposition, are significantly reduced in the vegetative and central cells in both eudicots and monocot species (Baroux and Autran, 2015). The chromatin decondensation in companion cells seems to influence the epigenetic setup of the gamete cells. In the companion cells, massive transcription of TEs occurs (following active epigenetic marks) and then TE-specific siRNAs are generated as a consequence (Schoft et al., 2009; Slotkin et al., 2009). These siRNAs can travel into the corresponding gamete cell and act *in-trans* on the chromatin to reinforce TE silencing by inducing the RNA-directed DNA methylation pathway, indicative of the importance of companion cells in maintaining genome stability and integrity of the gametes (Ibarra et al., 2012). *Trans*-silencing of a reporter gene GFP was successfully achieved in the sperm cells by expressing a corresponding amiRNA in the vegetative cell. This result supports an idea that siRNA can move from the companion cell to the male gametes, and the mobility of siRNAs was also confirmed between the central cell and the egg cell (Ibarra et al., 2012; Feng et al., 2013).

In contrast to the companion cell chromatin, the gamete chromatin exhibits more pronounced condensation. The highly deposited H3K9me2 modification of the sperm chromatin, especially at heterochromatic loci, partially contributes to the condensed chromatin state. The egg cell chromatin also harbors much more H3K9me2 (Pillot et al., 2010; She and Baroux, 2014). Moreover, other repressive epigenetic marks and associated enzymes are also enriched in the egg cell; therefore, the transcription in the egg cell is almost at a quiescent state, coincident with low-to-undetectable levels of the active RNA Pol II (Pillot et al., 2010). Paradoxically, the transcriptionally permissive mark H3K4me2 is enriched and the repressive mark

H3K27me3 is globally depleted in the sperm chromatin, which might be essential events for the transcription of sperm-specific genes (Pillot et al., 2010; She and Baroux, 2014). It is well-known that the transcriptomes of both the female and male gametes are characterized by a set of specifically expressed genes that are otherwise silenced in the somatic tissues. Thus, cell-specific epigenetic landscapes occurring during gametogenesis may create a favorable environment for the de-repression of those gamete-specific genes.

Reprogramming of H3 Lysine Methylation During Fertilization and Pre-embryogenesis

In angiosperms, two fertilization products are generated following double fertilization, a specific process in which two haploid sperm cells (*1n*) are delivered to the embryo sac through the pollen tube and simultaneously to fertilize with the haploid egg cell (*1n*) and the homodiploid central cell (*2n*) to generate a diploid embryo (*2n*) and a triploid endosperm *(3n*), respectively (reviewed in Bleckmann et al., 2014). The dimorphic epigenetic chromatin status in the egg cell and central cell directly gives rise to the dimorphism of the chromatin and transcriptional status in these two fertilized products, the embryo (zygote) and the endosperm (Pillot et al., 2010). The highly permissive and transcriptionally active state of the central cell is largely inherited by the endosperm following fertilization, therefore the chromatin dynamics in the central cell is likely to be a pre-patterning event for its post-fertilization fate.

Genomic imprinting is the consequence of the dimorphic epigenetic status in the asymmetric epigenetic setup between the embryo and the endosperm (Rodrigues and Zilberman, 2015), and imprinting regulation involves PRC2-mediated histone modification and likely other epigenetic mechanisms. For instance, genes with permissive epigenetic marks in the central cell, but are highly repressed in the condensed sperm chromatin, will develop into maternally expressed imprinted genes (MEGs) in the endosperm after fertilization. The MEGs and the paternally-expressed imprinted genes (PEGs) associate closely with differentially methylated regions (DMRs) between the paternal and the maternal genome (Zhang et al., 2014; Wang et al., 2015; Wang and Kohler, 2017). The PRC2 complex can target these DMRs for H3K27me3, and it has been reported that H3K27me3 can be deposited at hypomethylated regions in the maternal genome to determine the imprinted expression of PEGs, indicating that paternally and maternally hypomethylated regions contribute to the silencing of neighboring genes (Hsieh et al., 2011; Wolff et al., 2011; Borg et al., 2020). Additionally, several MEGs in the endosperm, such as MEA and FIS2, are essential for seed development, and mutations in *MEA* and *FIS2* cause seed abortion after fertilization. A subset of PEGs has indeed been demonstrated to have functions in building interploidy hybridization barriers in *Arabidopsis* (Wolff et al., 2011; Kradolfer et al., 2013).

The pre-embryo seems in a quiescent transcriptional state with a barely detectable Pol II activity, while the endosperm harbors a transcriptionally active chromatin status as shown by abundant levels of engaged RNA Pol II (Pillot et al., 2010). The distinct chromatin and transcriptional states of the fertilized products are largely inherited from their female gametic progenitors, the egg cell or the central cell, thus H3K9me2 is enriched in the zygote but reduced in the endosperm. Similar to the situation in the gametes and their companion cells, the transcription of TEs in the endosperm is derepressed to produce TE specific siRNAs that travel into the embryo to reinforce TE silencing in the zygote (Mosher et al., 2009; Ibarra et al., 2012). The embryo has the potency to develop into a future plant to establish novel cell types and organ symmetries; therefore, the newly formed zygote must be released from the gametic programs to obtain totipotency. The rapid reprogramming of chromatin status and histone medication in the zygote might be required for the establishment of future totipotency.

DISCUSSION AND PERSPECTIVES

Chromatin reprogramming during gametogenesis, fertilization, and early embryonic development is crucial not only in maintaining genomic integrity but also in setting pluri-or totipotency and resetting silenced genes necessary for the plant life cycle. Although some evidence has shown the existence of H3 lysine reprogramming during plant sexual reproduction, especially during cell fate specification, limitations in cell-specific epigenomic techniques still leave this exciting problem in a state of incomprehension. Great efforts are still required to overcome obstacles in cell-specific epigenomic profiling of the reproductive lineage, particularly in the model plant *Arabidopsis thaliana* with extremely small germ cells, zygote, and endosperm. The development of cell-specific nuclei isolation approaches, including INTACT (isolation of nuclei tagged in specific cell types) (Deal and Henikoff, 2011; Moreno-Romero et al., 2017) and FACS (fluorescence-activated cell sorting) (Moreno-Romero et al., 2017; Gustafsson et al., 2019; Nott et al., 2021), may prove to be a real asset in these efforts, though it still requires improvement in optimization. Single cell epigenomics is an inevitable solution to reveal the chromatin remodeling in multiple different cell types in the reproductive lineage, and to provide more accurate and integrated interpretation to fully understand the role of reprogramming events in functional gamete formation and seed development.

Epigenetic modification is a reversible mark, which can be removed from or redeposited to target genes to affect their expression. It will be of great interest to learn where, when, and how histone modification reprogramming occurs to reset the expression of those genes in different generations, two characteristic examples being the resetting of *FLOWERING LOCUS C* (*FLC*) (Sheldon et al., 2008) and miR156/157, the master regulator of vegetative phase change in plants (Nodine and Bartel, 2010). *FLC* expression is repressed mainly by VRN2-PRC2 mediated H3K27me3 deposition under vernalization or cold treatment until it is reset to an active transcriptional state during plant reproductive lineages. This off-reset process is mediated by depositing active epigenetic marks and by removing repressive marks (De Lucia et al., 2008; Sheldon et al., 2008;

Liu et al., 2021). In the pro-embryo stage, a seed-specific pioneer transcription factor, the *LEAFY COTYLEDON1* (*LEC1*), has been shown to establish active chromatin modifications at the *FLC* locus to re-trigger *FLC* expression (Tao et al., 2017). In early embryogenesis, two homologous B3 domain transcription factors LEC2 and FUSCA3 (FUS3) compete against two repressive modifiers to disrupt *FLC* silencing (Tao et al., 2019). These results suggest that the mechanism of gene off-reset pattern can be revealed by identifying specific transcription factors or histone modification associated co-factors specifically expressed in distinct reproductive tissues, or by searching readers containing bivalent or multivalent histone mark recognizing domains. In addition, histone readers might also recruit or stabilize various transcription factors, chromatin remodeling complexes and other components of the transcriptional network at the chromatin level to ensure proper transcriptional outcomes. Therefore, identifying the reader proteins and related complex might provide insight into the mechanism of gene off-reset during plant reproduction. A typical off-on resetting pattern during the plant life cycle occurs in the regulation of miR156/157, which is highly expressed in the juvenile phase, but declines gradually in the adult phase. The temporal expression pattern of miR156/157 during vegetative development is shown to be a result of the removal of active epigenetic marks and the deposition of some repressive epigenetic marks (Wolff et al., 2015; Xu et al., 2016, 2018; Fouracre et al., 2021). It is reasonable to assume that the silenced miR156/miR157 should be re-activated in the gametogenesis or pre-embryo stage to maintain its higher expression level in the juvenile phase in the next generation.

The mechanism by which the cell receives the instruction during plant sexual reproduction to initiate histone modification reprogramming is still poorly understood, and how to finely regulate histone modification at specific time points in specific cells still remains an open scientific question. As various epigenetic events, especially histone modification and DNA methylation, cooperate and interplay with each other closely; therefore, revealing the reprogramming and initiation of DNA methylation and other epigenetic marks during sexual reproduction will provide meaningful references. In addition, a reasonable explanation is that specific histone modification associated enzymes are recruited by specific cofactors, binding effectors, or transcription factors, which might be exclusively expressed in special cell types or selectively expressed at a certain developmental point, to specific sites to initiate histone modification. Therefore, searching for the developmental stage-specific and cell type-specific transcription factors or associated co-factors for histone modifier recruitment to dynamically regulate the chromatin state at specific loci will be one of the major tasks in future research.

AUTHOR CONTRIBUTIONS

HF conceived the present idea and wrote the manuscript. HF and YS consulted and collected relevant references. GW supervised the project, provided critical feedback, and helped shape the final manuscript. GW and YS helped revise and proofread the manuscript. All authors discussed the results and commented on the manuscript.

FUNDING

This work was supported by a start-up fund from Zhejiang Agricultural and Forestry University (Grant No. 2020FR035), and a fund from the Natural Science Foundation of Zhejiang Province (Grant No. Q21C060002) to HF. This work was also supported by the National Natural Science Foundation of China (31770209 and 31970191) to GW.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.782450/full#supplementary-material

REFERENCES

Alvarez-Venegas, R., and Avramova, Z. (2005). Methylation patterns of histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive *Arabidopsis* genes and in *atx1* mutants. *Nucleic Acids Res.* 33, 5199–5207. doi: 10.1093/nar/gki830

Alvarez-Venegas, R., Pien, S., Sadder, M., Witmer, X., Grossniklaus, U., et al. (2003). Atx-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. *Curr. Biol.* 13, 627–637. doi: 10.1016/s0960-9822(03)00243-4

Atlasi, Y., and Stunnenberg, H. G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* 18, 643–658. doi: 10.1038/nrg.2017.57

Baile, F., Merini, W., Hidalgo, I., and Calonje, M. (2021). EAR domain-containing transcription factors trigger PRC2-mediated chromatin marking in arabidopsis. *Plant Cell* 33, 2701–2715. doi: 10.1093/plcell/koab139

Bao, J., and Yan, W. (2012). Male germline control of transposable elements. *Biol. Reprod.* 86, 1–14. doi: 10.1095/biolreprod.111.095463

Baroux, C., and Autran, D. (2015). Chromatin dynamics during cellular differentiation in the female reproductive lineage of flowering plants. *Plant J.* 83, 160–176. doi: 10.1111/tpj.12890

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

Berger, F., and Twell, D. (2011). Germline specification and function in plants. *Annu. Rev. Plant Biol.* 62, 461–484. doi: 10.1146/annurev-arplant-042110-103824

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature* 447, 407–412. doi: 10.1038/nature05915

Berr, A., McCallum, E. J., Menard, R., Meyer, D., Fuchs, J., et al. (2010). *Arabidopsis SET DOMAIN GROUP2* is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell* 22, 3232–3248. doi: 10.1105/tpc.110.079962

Berr, A., Shafiq, S., Pinon, V., Dong, A., and Shen, W. H. (2015). The trxG family histone methyltransferase SET DOMAIN GROUP 26 promotes flowering via a distinctive genetic pathway. *Plant J.* 81, 316–328. doi: 10.1111/tpj.12729

Berr, A., Xu, L., Gao, J., Cognat, V., Steinmetz, A., et al. (2009). *SET DOMAIN GROUP25* encodes a histone methyltransferase and is involved in *FLOWERING LOCUS C* activation and repression of flowering. *Plant Physiol.* 151, 1476–1485. doi: 10.1104/pp.109.143941

Bleckmann, A., Alter, S., and Dresselhaus, T. (2014). The beginning of a seed: regulatory mechanisms of double fertilization. *Front. Plant Sci.* 5:452. doi: 10.3389/fpls.2014.00452

Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendia, D., et al. (2020). Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat. Cell Biol.* 22, 621–629. doi: 10.1038/s41556-020-0515-y

Caro, E., Stroud, H., Greenberg, M. V., Bernatavichute, Y. V., Feng, S., et al. (2012). The SET-domain protein SUVR5 mediates H3K9me2 deposition and silencing at stimulus response genes in a DNA methylation-independent manner. *PLoS Genet.* 8:e1002995. doi: 10.1371/journal.pgen.1002995

Cartagena, J. A., Matsunaga, S., Seki, M., Kurihara, D., Yokoyama, M., et al. (2008). The *Arabidopsis* SDG4 contributes to the regulation of pollen tube growth by methylation of histone H3 lysines 4 and 36 in mature pollen. *Dev. Biol.* 315, 355–368. doi: 10.1016/j.ydbio.2007.12.016

Cazzonelli, C. I., Cuttriss, A. J., Cossetto, S. B., Pye, W., Crisp, P., et al. (2009). Regulation of carotenoid composition and shoot branching in *Arabidopsis* by a chromatin modifying histone methyltransferase, SDG8. *Plant Cell* 21, 39–53. doi: 10.1105/tpc.108.063131

Charron, J. B., He, H., Elling, A. A., and Deng, X. W. (2009). Dynamic landscapes of four histone modifications during deetiolation in *Arabidopsis*. *Plant Cell* 21, 3732–3748. doi: 10.1105/tpc.109.066845

Chen, C., Farmer, A. D., Langley, R. J., Mudge, J., Crow, J. A., et al. (2010). Meiosis-specific gene discovery in plants: RNA-seq applied to isolated *Arabidopsis* male meiocytes. *BMC Plant Biol.* 10:280. doi: 10.1186/1471-2229-10-280

Chen, L. Q., Luo, J. H., Cui, Z. H., Xue, M., Wang, L., et al. (2017). *ATX3*, *ATX4*, and *ATX5* encode putative H3K4 methyltransferases and are critical for plant development. *Plant Physiol.* 174, 1795–1806. doi: 10.1104/pp.16.01944

Cheng, K., Xu, Y., Yang, C., Ouellette, L., Niu, L., et al. (2020). Histone tales: lysine methylation, a protagonist in arabidopsis development. *J. Exp. Bot.* 71, 793–807. doi: 10.1093/jxb/erz435

Crevillen, P. (2020). Histone demethylases as counterbalance to H3K27me3 silencing in plants. *Iscience* 23:101715. doi: 10.1016/j.isci.2020.101715

Dahia, P., Clifton-Bligh, R., Gimenez-Roqueplo, A. P., Robledo, M., and Jimenez, C. (2020). Hereditary endocrine tumours: current state-of-the-art and research opportunities: metastatic pheochromocytomas and paragangliomas: proceedings of the MEN2019 workshop. *Endocr. Relat. Cancer* 27, T41–T52. doi: 10.1530/ERC-19-0435

de la Paz, S. M., Aceves-Garcia, P., Petrone, E., Steckenborn, S., Vega-Leon, R., et al. (2015). The impact of polycomb group (PcG) and trithorax group (TrxG) epigenetic factors in plant plasticity. *New Phytol.* 208, 684–694. doi: 10.1111/nph.13486

de la Paz, S. M., and Gutierrez, C. (2009). *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. *Proc. Natl. Acad. Sci. U S A* 106, 2065–2070. doi: 10.1073/pnas.0811093106

De Lucia, F., Crevillen, P., Jones, A. M., Greb, T., and Dean, C. (2008). A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc. Natl. Acad. Sci. U S A* 105, 16831–16836. doi: 10.1073/pnas.0808687105

Deal, R. B., and Henikoff, S. (2011). The intact method for cell type-specific gene expression and chromatin profiling in *Arabidopsis thaliana*. *Nat. Protoc.* 6, 56–68. doi: 10.1038/nprot.2010.175

Ding, Y., Avramova, Z., and Fromm, M. (2011). Two distinct roles of ARABIDOPSIS HOMOLOG OF TRITHORAX1 (ATX1) at promoters and within transcribed regions of ATX1-regulated genes. *Plant Cell* 23, 350–363. doi: 10.1105/tpc.110.080150

Dong, G., Ma, D. P., and Li, J. (2008). The histone methyltransferase SDG8 regulates shoot branching in *Arabidopsis*. *Biochem. Biophys. Res. Commun.* 373, 659–664. doi: 10.1016/j.bbrc.2008.06.096

Drews, G. N., and Koltunow, A. M. (2011). The female gametophyte. *Arabidopsis Book* 9:e155. doi: 10.1199/tab.0155

Du, J., Johnson, L. M., Groth, M., Feng, S., Hale, C. J., et al. (2014). Mechanism of DNA methylation-directed histone methylation by kryptonite. *Mol. Cell* 55, 495–504. doi: 10.1016/j.molcel.2014.06.009

Dutta, A., Choudhary, P., Caruana, J., and Raina, R. (2017). JMJ27, an Arabidopsis H3K9 histone demethylase, modulates defense against *Pseudomonas syringae* and flowering time. *Plant J.* 91, 1015–1028. doi: 10.1111/tpj.13623

Ebbs, M. L., and Bender, J. (2006). Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. *Plant Cell* 18, 1166–1176. doi: 10.1105/tpc.106.041400

Ebbs, M. L., Bartee, L., and Bender, J. (2005). H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol. Cell. Biol.* 25, 10507–10515. doi: 10.1128/MCB.25.23.10507-10515.2005

Feng, S., and Jacobsen, S. E. (2011). Epigenetic modifications in plants: an evolutionary perspective. *Curr. Opin. Plant Biol.* 14, 179–186. doi: 10.1016/j.pbi.2010.12.002

Feng, S., Jacobsen, S. E., and Reik, W. (2010). Epigenetic reprogramming in plant and animal development. *Science* 330, 622–627. doi: 10.1126/science.1190614

Feng, X., Zilberman, D., and Dickinson, H. (2013). A conversation across generations: soma-germ cell crosstalk in plants. *Dev. Cell* 24, 215–225. doi: 10.1016/j.devcel.2013.01.014

Fiorucci, A. S., Bourbousse, C., Concia, L., Rougee, M., Deton-Cabanillas, A. F., et al. (2019). Arabidopsis S2Lb links ATCOMPASS-like and SDG2 activity in H3K4me3 independently from histone H2B monoubiquitination. *Genome Biol.* 20:100. doi: 10.1186/s13059-019-1705-4

Foroozani, M., Vandal, M. P., and Smith, A. P. (2021). H3K4 trimethylation dynamics impact diverse developmental and environmental responses in plants. *Planta* 253:4. doi: 10.1007/s00425-020-03520-0

Fouracre, J. P., He, J., Chen, V. J., Sidoli, S., and Poethig, R. S. (2021). *VAL* genes regulate vegetative phase change via miR156-dependent and independent mechanisms. *PLoS Genet.* 17:e1009626. doi: 10.1371/journal.pgen.1009626

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns–from conservation to diversity. *Trends Plant Sci.* 11, 199–208. doi: 10.1016/j.tplants.2006.02.008

Gan, E. S., Xu, Y., Wong, J. Y., Goh, J. G., Sun, B., et al. (2014). Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of *FLC* in *Arabidopsis*. *Nat. Commun.* 5:5098. doi: 10.1038/ncomms6098

Gehring, M. (2019). Epigenetic dynamics during flowering plant reproduction: evidence for reprogramming? *New Phytol.* 224, 91–96. doi: 10.1111/nph.15856

Gehring, M., and Satyaki, P. R. (2017). Endosperm and imprinting, inextricably linked. *Plant Physiol.* 173, 143–154. doi: 10.1104/pp.16.01353

Golbabapour, S., Majid, N. A., Hassandarvish, P., Hajrezaie, M., Abdulla, M. A., et al. (2013). Gene silencing and polycomb group proteins: an overview of their structure, mechanisms and phylogenetics. *Omics* 17, 283–296. doi: 10.1089/omi.2012.0105

Greer, E. L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* 13, 343–357. doi: 10.1038/nrg3173

Grini, P. E., Thorstensen, T., Alm, V., Vizcay-Barrena, G., Windju, S. S., et al. (2009). The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in *Arabidopsis*. *PLoS One* 4:e7817. doi: 10.1371/journal.pone.0007817

Guo, L., Yu, Y., Law, J. A., and Zhang, X. (2010). SET DOMAIN GROUP2 is the major histone H3 lysie 4 trimethyltransferase in *Arabidopsis* (vol 107, pg 18557, 2010). *Proc. Natl. Acad. Sci.* 107, 18557–18562.

Gustafsson, C., De Paepe, A., Schmidl, C., and Mansson, R. (2019). High-throughput chipmentation: freely scalable, single day chipseq data generation from very low cell-numbers. *BMC Genomics* 20:59. doi: 10.1186/s12864-018-5299-0

Gutierrez-Marcos, J. F., and Dickinson, H. G. (2012). Epigenetic reprogramming in plant reproductive lineages. *Plant Cell Physiol.* 53, 817–823. doi: 10.1093/pcp/pcs052

Han, Q., Bartels, A., Cheng, X., Meyer, A., An, Y. C., et al. (2019). Epigenetics regulates reproductive development in plants. *Plants* 8:lants8120564. doi: 10.3390/plants8120564

Hiragami-Hamada, K., Soeroes, S., Nikolov, M., Wilkins, B., Kreuz, S., et al. (2016). Dynamic and flexible H3K9me3 bridging via hp1beta dimerization establishes a plastic state of condensed chromatin. *Nat. Commun.* 7:11310. doi: 10.1038/ncomms11310

Hofmann, F., Schon, M. A., and Nodine, M. D. (2019). The embryonic transcriptome of *Arabidopsis thaliana*. *Plant Reprod.* 32, 77–91. doi: 10.1007/s00497-018-00357-2

Hsieh, T. F., Shin, J., Uzawa, R., Silva, P., Cohen, S., et al. (2011). Regulation of imprinted gene expression in *Arabidopsis* endosperm. *Proc. Natl. Acad. Sci. U S A* 108, 1755–1762. doi: 10.1073/pnas.1019273108

Hu, H., Tian, S., Xie, G., Liu, R., Wang, N., et al. (2021). TEM1 combinatorially binds to *FLOWERING LOCUS T* and recruits a polycomb factor to repress the floral transition in *Arabidopsis*. *Proc. Natl. Acad. Sci. U S A* 35:2103895118. doi: 10.1073/pnas.2103895118

Ibarra, C. A., Feng, X., Schoft, V. K., Hsieh, T. F., Uzawa, R., et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* 337, 1360–1364. doi: 10.1126/science.1224839

Jackson, J. P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., et al. (2004). Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in *Arabidopsis thaliana*. *Chromosoma* 112, 308–315. doi: 10.1007/s00412-004-0275-7

Jackson, J. P., Lindroth, A. M., Cao, X., and Jacobsen, S. E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556–560. doi: 10.1038/nature731

Jacob, Y., Feng, S., LeBlanc, C. A., Bernatavichute, Y. V., Stroud, H., et al. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat. Struct. Mol. Biol.* 16, 763–768. doi: 10.1038/nsmb.1611

Jiang, D., Gu, X., He, et al. (2009). Establishment of the winter-annual growth habit via *FRIGIDA*-mediated histone methylation at *FLOWERING LOCUS C* in *Arabidopsis*. *Plant Cell* 21, 1733–1746.

Jiang, D., Yang, W., He, Y., and Amasino, R. M. (2007). *Arabidopsis* relatives of the human lysine-specific demethylase1 repress the expression of *FWA* and *FLOWERING LOCUS C* and thus promote the floral transition. *Plant Cell* 19, 2975–2987. doi: 10.1105/tpc.107.052373

Johnson, L., Mollah, S., Garcia, B. A., Muratore, T. L., Shabanowitz, J., et al. (2004). Mass spectrometry analysis of *Arabidopsis* histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res.* 32, 6511–6518. doi: 10.1093/nar/gkh992

Kahn, T. G., Dorafshan, E., Schultheis, D., Zare, A., Stenberg, P., et al. (2016). Interdependence of PRC1 and PRC2 for recruitment to polycomb response elements. *Nucleic Acids Res.* 44, 10132–10149. doi: 10.1093/nar/gkw701

Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. *Nat. Rev. Genet.* 15, 613–624. doi: 10.1038/nrg3685

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., and Penin, A. A. (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.* 88, 1058–1070. doi: 10.1111/tpj.13312

Klose, R. J., and Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. *Nat. Rev. Mol. Cell Biol.* 8, 307–318. doi: 10.1038/nrm2143

Kohler, C., and Villar, C. B. (2008). Programming of gene expression by Polycomb group proteins. *Trends Cell Biol.* 18, 236–243. doi: 10.1016/j.tcb.2008.02.005

Kohler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U., et al. (2003). *Arabidopsis* MSI1 is a component of the MEA/FIE *Polycomb* group complex and required for seed development. *EMBO J.* 22, 4804–4814. doi: 10.1093/emboj/cdg444

Kradolfer, D., Wolff, P., Jiang, H., Siretskiy, A., and Kohler, C. (2013). An imprinted gene underlies postzygotic reproductive isolation in *Arabidopsis thaliana*. *Dev. Cell* 26, 525–535. doi: 10.1016/j.devcel.2013.08.006

Kuwano, M., Masumura, T., and Yoshida, K. T. (2011). A novel endosperm transfer cell-containing region-specific gene and its promoter in rice. *Plant Mol. Biol.* 76, 47–56. doi: 10.1007/s11103-011-9765-1

Lee, K., Park, O. S., and Seo, P. J. (2017). *Arabidopsis* ATXR2 deposits H3K36me3 at the promoters of *LBD* genes to facilitate cellular dedifferentiation. *Sci. Signal.* 10:aan0316. doi: 10.1126/scisignal.aan0316

Lee, K., Park, O. S., and Seo, P. J. (2018). JMJ30-mediated demethylation of H3K9me3 drives tissue identity changes to promote callus formation in Arabidopsis. *Plant J.* 95, 961–975. doi: 10.1111/tpj.14002

Li, C., Gu, L., Gao, L., Chen, C., Wei, C. Q., et al. (2016). Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in *Arabidopsis*. *Nat. Genet.* 48, 687–693. doi: 10.1038/ng.3555

Li, Y., Mukherjee, I., Thum, K. E., Tanurdzic, M., Katari, M. S., et al. (2015). The histone methyltransferase SDG8 mediates the epigenetic modification of light and carbon responsive genes in plants. *Genome Biol.* 16:79. doi: 10.1186/s13059-015-0640-2

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. *Annu. Rev. Plant Biol.* 61, 395–420. doi: 10.1146/annurev.arplant.043008.091939

Liu, F., Quesada, V., Crevillen, P., Baurle, I., Swiezewski, S., et al. (2007). The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate *FLC*. *Mol. Cell* 28, 398–407. doi: 10.1016/j.molcel.2007.10.018

Liu, P., Zhang, S., Zhou, B., Luo, X., Zhou, X. F., et al. (2019). The histone H3K4 demethylase JMJ16 represses leaf senescence in Arabidopsis. *Plant Cell* 31, 430–443. doi: 10.1105/tpc.18.00693

Liu, X., Luo, J., Li, T., Yang, H., Wang, P., et al. (2021). *SDG711* is involved in rice seed development through regulation of starch metabolism gene expression in coordination with other histone modifications. *Rice* 14:25. doi: 10.1186/s12284-021-00467-y

Liu, Y., and Huang, Y. (2018). Uncovering the mechanistic basis for specific recognition of monomethylated H3K4 by the CW domain of *Arabidopsis* histone methyltransferase SDG8. *J. Biol. Chem.* 293, 6470–6481. doi: 10.1074/jbc.RA117.001390

Liu, Y., Liu, K., Yin, L., Yu, Y., Qi, J., et al. (2019). H3K4me2 functions as a repressive epigenetic mark in plants. *Epigenet. Chromat.* 12:40. doi: 10.1186/s13072-019-0285-6

Liu, Z. W., Shao, C. R., Zhang, C. J., Zhou, J. X., Zhang, S. W., et al. (2014). The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. *PLoS Genet.* 10:e1003948. doi: 10.1371/journal.pgen.1003948

Lopato, S., Borisjuk, N., Langridge, P., and Hrmova, M. (2014). Endosperm transfer cell-specific genes and proteins: structure, function and applications in biotechnology. *Front. Plant Sci.* 5:64. doi: 10.3389/fpls.2014.00064

Loraine, A. E., McCormick, S., Estrada, A., Patel, K., and Qin, P. (2013). RNA-seq of Arabidopsis pollen uncovers novel transcription and alternative splicing. *Plant Physiol.* 162, 1092–1109. doi: 10.1104/pp.112.211441

Lu, F., Cui, X., Zhang, S., Liu, C., and Cao, X. (2010). JMJ14 is an H3K4 demethylase regulating flowering time in *Arabidopsis*. *Cell Res.* 20, 387–390. doi: 10.1038/cr.2010.27

Lu, F., Li, G., Cui, X., Liu, C., Wang, X. J., et al. (2008). Comparative analysis of jmjC domain-containing proteins reveals the potential histone demethylases in *Arabidopsis* and rice. *J. Integr. Plant Biol.* 50, 886–896. doi: 10.1111/j.1744-7909.2008.00692x

Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13, 436–447. doi: 10.1038/nrm3382

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389, 251–260. doi: 10.1038/38444

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J., and Chaudhury, A. (2000). Expression and parent-of-origin effects for *FIS2*, *MEA*, and *FIE* in the endosperm and embryo of developing *Arabidopsis* seeds. *Proc. Natl. Acad. Sci. U S A* 97, 10637–10642. doi: 10.1073/pnas.170292997

Macdonald, W. A. (2012). Epigenetic mechanisms of genomic imprinting: common themes in the regulation of imprinted regions in mammals, plants, and insects. *Genet. Res. Int.* 2012:585024. doi: 10.1155/2012/585024

Martignago, D., Bernardini, B., Polticelli, F., Salvi, D., Cona, A., et al. (2019). The four FAD-dependent histone demethylases of arabidopsis are differently involved in the control of flowering time. *Front. Plant Sci.* 10:669. doi: 10.3389/fpls.2019.00669

Mathieu, O., Probst, A. V., and Paszkowski, J. (2005). Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *EMBO J.* 24, 2783–2791. doi: 10.1038/sj.emboj.7600743

McLaughlin, N., Wang, F., Saifudeen, Z., and El-Dahr, S. S. (2014). In situ histone landscape of nephrogenesis. *Epigenetics* 9, 222–235. doi: 10.4161/epi.26793

Milutinovic, M., Lindsey, B. R., Wijeratne, A., Hernandez, J. M., Grotewold, N., et al. (2019). Arabidopsis EMSY-like (EML) histone readers are necessary for post-fertilization seed development, but prevent fertilization-independent seed formation. *Plant Sci.* 285, 99–109. doi: 10.1016/j.plantsci.2019.04.007

Miura, A., Nakamura, M., Inagaki, S., Kobayashi, A., Saze, H., et al. (2009). An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA

methylation at CHG sites. *EMBO J.* 28, 1078–1086. doi: 10.1038/emboj.2009.59

Molitor, A. M., Bu, Z., Yu, Y., and Shen, W. H. (2014). *Arabidopsis* AL PHD-PRC1 complexes promote seed germination through H3K4me3-to-H3K27me3 chromatin state switch in repression of seed developmental genes. *PLoS Genet.* 10:e1004091. doi: 10.1371/journal.pgen.1004091

Moreno-Romero, J., Santos-Gonzalez, J., Hennig, L., and Kohler, C. (2017). Applying the intact method to purify endosperm nuclei and to generate parental-specific epigenome profiles. *Nat. Protoc.* 12, 238–254. doi: 10.1038/nprot.2016.167

Mosher, R. A., Melnyk, C. W., Kelly, K. A., Dunn, R. M., Studholme, D. J., et al. (2009). Uniparental expression of Pol IV-dependent siRNAs in developing endosperm of *Arabidopsis*. *Nature* 460, 283–286. doi: 10.1038/nature08084

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., et al. (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J.* 24, 1418–1429. doi: 10.1038/sj.emboj.7600604

Ng, D. W., Wang, T., Chandrasekharan, M. B., Aramayo, R., Kertbundit, S., et al. (2007). Plant SET domain-containing proteins: structure, function and regulation. *Biochim. Biophys. Acta* 1769, 316–329. doi: 10.1016/j.bbaexp.2007.04.003

Niu, Q., Song, Z., Tang, K., Chen, L., and Lang, Z. (2021). A histone H3K4me1-specific binding protein is required for siRNA accumulation and DNA methylation at a subset of loci targeted by RNA-directed DNA methylation. *Nat. Commun.* 12:3367.

Nodine, M. D., and Bartel, D. P. (2010). MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev.* 24, 2678–2692. doi: 10.1101/gad.1986710

Nott, A., Schlachetzki, J., Fixsen, B. R., and Glass, C. K. (2021). Nuclei isolation of multiple brain cell types for omics interrogation. *Nat. Protoc.* 16, 1629–1646. doi: 10.1038/s41596-020-00472-3

Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., et al. (2010). Control of female gamete formation by a small RNA pathway in Arabidopsis. *Nature* 464, 628–632. doi: 10.1038/nature08828

Ono, A., and Kinoshita, T. (2021). Epigenetics and plant reproduction: Multiple steps for responsibly handling succession. *Curr. Opin. Plant Biol.* 61:102032. doi: 10.1016/j.pbi.2021.102032

Pignatta, D., Novitzky, K., Satyaki, P., and Gehring, M. (2018). A variably imprinted epiallele impacts seed development. *PLoS Genet.* 14:e1007469. doi: 10.1371/journal.pgen.1007469

Pillot, M., Baroux, C., Vazquez, M. A., Autran, D., Leblanc, O., et al. (2010). Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* 22, 307–320. doi: 10.1105/tpc.109.071647

Pinon, V., Yao, X., Dong, A., and Shen, W. H. (2017). SDG2-mediated H3K4me3 is crucial for chromatin condensation and mitotic division during male gametogenesis in Arabidopsis. *Plant Physiol.* 174, 1205–1215. doi: 10.1104/pp.17.00306

Pontvianne, F., Blevins, T., and Pikaard, C. S. (2010). Arabidopsis histone lysine methyltransferases. *Adv. Bot. Res.* 53, 1–22. doi: 10.1016/S0065-2296(10)53001-5

Qian, S., Lv, X., Scheid, R. N., Lu, L., Yang, Z., et al. (2018). Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader shl. *Nat. Commun.* 9:2425. doi: 10.1038/s41467-018-04836-y

Rahmati, I. M., Brown, E., Weigand, C., Tillett, R. L., Schlauch, K. A., et al. (2018). A comparison of heat-stress transcriptome changes between wild-type *Arabidopsis* pollen and a heat-sensitive mutant harboring a knockout of *cyclic nucleotide-gated cation channel 16* (*cngc16*). *BMC Genomics* 19:549. doi: 10.1186/s12864-018-4930-4

Rodrigues, J. A., and Zilberman, D. (2015). Evolution and function of genomic imprinting in plants. *Genes Dev.* 29, 2517–2531. doi: 10.1101/gad.269902.115

Roudier, F., Teixeira, F. K., and Colot, V. (2009). Chromatin indexing in *Arabidopsis*: an epigenomic tale of tails and more. *Trends Genet.* 25, 511–517. doi: 10.1016/j.tig.2009.09.013

Ryu, H. Y., and Hochstrasser, M. (2021). Histone sumoylation and chromatin dynamics. *Nucleic Acids Res.* 49, 6043–6052. doi: 10.1093/nar/gkab280

Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., et al. (2008). The highly similar *Arabidopsis* homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell* 20, 568–579. doi: 10.1105/tpc.107.056614

Samo, N., Ebert, A., Kopka, J., and Mozgova, I. (2021). Plant chromatin, metabolism and development - an intricate crosstalk. *Curr. Opin. Plant Biol.* 61:102002. doi: 10.1016/j.pbi.2021.102002

Saze, H., Shiraishi, A., Miura, A., and Kakutani, T. (2008). Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* 319, 462–465. doi: 10.1126/science.1150987

Schmidt, A., Wohrmann, H. J., Raissig, M. T., Arand, J., Gheyselinck, J., et al. (2013). The *Polycomb* group protein MEDEA and the DNA methyltransferase MET1 interact to repress autonomous endosperm development in Arabidopsis. *Plant J.* 73, 776–787. doi: 10.1111/tpj.12070

Schoft, V. K., Chumak, N., Mosiolek, M., Slusarz, L., Komnenovic, V., et al. (2009). Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Rep.* 10, 1015–1021. doi: 10.1038/embor.2009.152

Shafiq, S., Berr, A., and Shen, W. H. (2014). Combinatorial functions of diverse histone methylations in arabidopsis thaliana flowering time regulation. *New Phytol.* 201, 312–322. doi: 10.1111/nph.12493

She, W., and Baroux, C. (2014). Chromatin dynamics during plant sexual reproduction. *Front. Plant Sci.* 5:354. doi: 10.3389/fpls.2014.00354

She, W., and Baroux, C. (2015). Chromatin dynamics in pollen mother cells underpin a common scenario at the somatic-to-reproductive fate transition of both the male and female lineages in *Arabidopsis*. *Front. Plant Sci.* 6:294. doi: 10.3389/fpls.2015.00294

She, W., Grimanelli, D., Rutowicz, K., Whitehead, M. W., Puzio, M., et al. (2013). Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. *Development* 140, 4008–4019. doi: 10.1242/dev.095034

Sheldon, C. C., Hills, M. J., Lister, C., Dean, C., Dennis, E. S., et al. (2008). Resetting of *FLOWERING LOCUS C* expression after epigenetic repression by vernalization. *Proc. Natl. Acad. Sci. U S A* 105, 2214–2219. doi: 10.1073/pnas.0711453105

Slotkin, R. K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J. D., et al. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136, 461–472. doi: 10.1016/j.cell.2008.12.038

Springer, N. M., Napoli, C. A., Selinger, D. A., Pandey, R., Cone, K. C., et al. (2003). Comparative analysis of SET domain proteins in Maize and Arabidopsis reveals multiple duplications preceding the divergence of Monocots and Dicots. *Plant Physiol.* 132, 907–925.

Tamada, Y., Yun, J. Y., Woo, S. C., and Amasino, R. M. (2009). ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of *FLOWERING LOCUS C*. *Plant Cell* 21, 3257–3269. doi: 10.1105/tpc.109.070060

Tao, Z., Hu, H., Luo, X., Jia, B., Du, J., et al. (2019). Embryonic resetting of the parental vernalized state by two B3 domain transcription factors in *Arabidopsis*. *Nat. Plants* 5, 424–435. doi: 10.1038/s41477-019-0402-3

Tao, Z., Shen, L., Gu, X., Wang, Y., Yu, H., et al. (2017). Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature* 551, 124–128. doi: 10.1038/nature24300

Tedeschi, F., Rizzo, P., Rutten, T., Altschmied, L., and Baumlein, H. (2017). RWP-RK domain-containing transcription factors control cell differentiation during female gametophyte development in Arabidopsis. *New Phytol.* 213, 1909–1924. doi: 10.1111/nph.14293

Thorstensen, T., Fischer, A., Sandvik, S. V., Johnsen, S. S., Grini, P. E., et al. (2006). The *Arabidopsis* SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic Acids Res.* 34, 5461–5470. doi: 10.1093/nar/gkl687

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M. L., Guillaume, E., et al. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* 3:e86. doi: 10.1371/journal.pgen.0030086

Twell, D. (2011). Male gametogenesis and germline specification in flowering plants. *Sex. Plant Reprod.* 24, 149–160. doi: 10.1007/s00497-010-0157-5

Wang, G., and Kohler, C. (2017). Epigenetic processes in flowering plant reproduction. *J. Exp. Bot.* 68, 797–807. doi: 10.1093/jxb/erw486

Wang, P., Xia, H., Zhang, Y., Zhao, S., Zhao, C., et al. (2015). Genome-wide high-resolution mapping of DNA methylation identifies epigenetic variation

across embryo and endosperm in Maize (*Zea may*). *BMC Genomics* 16:21. doi: 10.1186/s12864-014-1204-7

Wang, Y., and Copenhaver, G. P. (2018). Meiotic recombination: Mixing it up in plants. *Annu. Rev. Plant Biol.* 69, 577–609. doi: 10.1146/annurev-arplant-042817-040431

Wang, Y., Zhou, X., Luo, J., Lv, S., Liu, R., et al. (2021). Recognition of H3K9me1 by maize RNA-directed DNA methylation factor SHH2. *J. Integr. Plant Biol.* 63, 1091–1096. doi: 10.1111/jipb.13103

Weinhofer, I., Hehenberger, E., Roszak, P., Hennig, L., and Kohler, C. (2010). H3K27me3 profiling of the endosperm implies exclusion of polycomb group protein targeting by DNA methylation. *PLoS Genet.* 6:1001152. doi: 10.1371/journal.pgen.1001152

Willmann, M. R., Berkowitz, N. D., and Gregory, B. D. (2014). Improved genome-wide mapping of uncapped and cleaved transcripts in eukaryotes–GMUCT 2.0. *Methods* 67, 64–73. doi: 10.1016/j.ymeth.2013.07.003

Wolff, P., Jiang, H., Wang, G., Santos-Gonzalez, J., and Kohler, C. (2015). Paternally expressed imprinted genes establish postzygotic hybridization barriers in *Arabidopsis thaliana*. *Elife* 4:10074. doi: 10.7554/eLife.10074

Wolff, P., Weinhofer, I., Seguin, J., Roszak, P., Beisel, C., et al. (2011). High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis endosperm. *PLoS Genet.* 7:e1002126. doi: 10.1371/journal.pgen.1002126

Xiao, J., and Wagner, D. (2015). Polycomb repression in the regulation of growth and development in *Arabidopsis*. *Curr. Opin. Plant Biol.* 23, 15–24. doi: 10.1016/j.pbi.2014.10.003

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb Repressive Complex 2 in Arabidopsis. *Nat. Genet.* 49, 1546–1552. doi: 10.1038/ng.3937

Xiao, J., Lee, U. S., and Wagner, D. (2016). Tug of war: adding and removing histone lysine methylation in *Arabidopsis*. *Curr. Opin. Plant Biol.* 34, 41–53. doi: 10.1016/j.pbi.2016.08.002

Xu, L., and Jiang, H. (2020). Writing and reading histone H3 lysine 9 methylation in *Arabidopsis*. *Front. Plant Sci.* 11:452. doi: 10.3389/fpls.2020.00452

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J. P., et al. (2008). Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Mol. Cell. Biol.* 28, 1348–1360. doi: 10.1128/MCB.01607-07

Xu, M., Hu, T., Smith, M. R., and Poethig, R. S. (2016). Epigenetic regulation of vegetative phase change in arabidopsis. *Plant Cell* 28, 28–41. doi: 10.1105/tpc.15.00854

Xu, M., Leichty, A. R., Hu, T., and Poethig, R. S. (2018). H2A.Z promotes the transcription of *MIR156A* and *MIR156C* in *Arabidopsis* by facilitating the deposition of H3K4me3. *Development* 145:152868. doi: 10.1242/dev.152868

Xu, Y., Gan, E. S., Zhou, J., Wee, W. Y., Zhang, X., et al. (2014). Arabidopsis MRG domain proteins bridge two histone modifications to elevate expression of flowering genes. *Nucleic Acids Res.* 42, 10960–10974. doi: 10.1093/nar/gku781

Yan, Y., Shen, L., Chen, Y., Bao, S., Thong, Z., et al. (2014). A MYB-domain protein EFM mediates flowering responses to environmental cues in *Arabidopsis*. *Dev. Cell* 30, 437–448. doi: 10.1016/j.devcel.2014.07.004

Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F., et al. (2013). VAL- and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in *Arabidopsis*. *Curr. Biol.* 23, 1324–1329. doi: 10.1016/j.cub.2013.05.050

Yang, H., Mo, H., Fan, D., Cao, Y., Cui, S., et al. (2012). Overexpression of a histone H3K4 demethylase, JMJ15, accelerates flowering time in *Arabidopsis*. *Plant Cell Rep.* 31, 1297–1308. doi: 10.1007/s00299-012-1249-5

Yang, Z., Qian, S., Scheid, R. N., Lu, L., Chen, X., et al. (2018). EBS is a bivalent histone reader that regulates floral phase transition in *Arabidopsis*. *Nat. Genet.* 50, 1247–1253. doi: 10.1038/s41588-018-0187-8

Yu, X., Li, L., Li, L., Guo, M., Chory, J., et al. (2008). Modulation of brassinosteroid-regulated gene expression by jumonji domain-containing proteins ELF6 and REF6 in *Arabidopsis*. *Proc. Natl. Acad. Sci. U S A* 105, 7618–7623. doi: 10.1073/pnas.0802254105

Yuan, L., Song, X., Zhang, L., Yu, Y., Liang, Z., et al. (2021). The transcriptional repressors VAL1 and VAL2 recruit PRC2 for genome-wide Polycomb silencing in *Arabidopsis*. *Nucleic Acids Res.* 49, 98–113. doi: 10.1093/nar/gkaa1129

Yun, J. Y., Tamada, Y., Kang, Y. E., and Amasino, R. M. (2012). Arabidopsis trithorax-related3/SET domain group2 is required for the winter-annual habit of *Arabidopsis thaliana*. *Plant Cell Physiol.* 53, 834–846. doi: 10.1093/pcp/pcs021

Zhang, H., Zhang, F., Yu, Y., Feng, L., Jia, J., et al. (2020). A comprehensive online database for exploring approximately 20,000 public *Arabidopsis* RNA-seq libraries. *Mol. Plant.* 13, 1231–1233. doi: 10.1016/j.molp.2020.08.001

Zhang, M., Xie, S., Dong, X., Zhao, X., Zeng, B., et al. (2014). Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. *Genome Res.* 24, 167–176. doi: 10.1101/gr.155879.113

Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., and Jacobsen, S. E. (2009). Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10:R62. doi: 10.1186/gb-2009-10-6-r62

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., et al. (2007a). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* 5:e129. doi: 10.1371/journal.pbio.0050129

Zhang, X., Germann, S., Blus, B. J., Khorasanizadeh, S., Gaudin, V., et al. (2007b). The *Arabidopsis* LHP1 protein colocalizes with histone H3 lys27 trimethylation. *Nat. Struct. Mol. Biol.* 14, 869–871. doi: 10.1038/nsmb1283

Zhang, X., Menard, R., Li, Y., Coruzzi, G. M., Heitz, T., et al. (2020). *Arabidopsis* SDG8 potentiates the sustainable transcriptional induction of the *Pathogenesis-Related* genes *PR1* and *PR2* during plant defense response. *Front. Plant Sci.* 11:277. doi: 10.3389/fpls.2020.00277

Zhao, S., Zhang, B., Yang, M., Zhu, J., and Li, H. (2018). Systematic profiling of histone readers in *Arabidopsis thaliana*. *Cell Rep.* 22, 1090–1102. doi: 10.1016/j.celrep.2017.12.099

Zhao, T., Zhan, Z., and Jiang, D. (2019). Histone modifications and their regulatory roles in plant development and environmental memory. *J. Genet. Genomics* 46, 467–476. doi: 10.1016/j.jgg.2019.09.005

Zheng, B., and Chen, X. (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant Biol.* 14, 123–129. doi: 10.1016/j.pbi.2011.01.001

Zheng, S., Hu, H., Ren, H., Yang, Z., Qiu, Q., et al. (2019). The Arabidopsis H3K27me3 demethylase JUMONJI 13 is a temperature and photoperiod dependent flowering repressor. *Nat. Commun.* 10:1303. doi: 10.1038/s41467-019-09310-x

Zhong, P., Li, J., Luo, L., Zhao, Z., and Tian, Z. (2019). Top1α regulates *FLOWERING LOCUS C* expression by coupling histone modification and transcription machinery. *Development* 146:167841.

Zhou, H., Liu, Y., Liang, Y., Zhou, D., Li, S., et al. (2020). The function of histone lysine methylation related SET domain group proteins in plants. *Protein Sci.* 29, 1120–1137. doi: 10.1002/pro.3849

Zhou, M., Palanca, A., and Law, J. A. (2018). Locus-specific control of the de novo DNA methylation pathway in *Arabidopsis* by the CLASSY family. *Nat. Genet.* 50, 865–873. doi: 10.1038/s41588-018-0115-y

Zhou, Y., Wang, Y., Krause, K., Yang, T., Dongus, J. A., et al. (2018). Telobox motifs recruit CLF/SWN-PRC2 for H3K27me3 deposition via TRB factors in *Arabidopsis*. *Nat. Genet.* 50, 638–644. doi: 10.1038/s41588-018-0109-9

Zy, Z. P., Liu, Y. C., Jiang, et al. (2014). Regulation of *Arabidopsis* flowering by the histone mark readers MRG1/2 via interaction with CONSTANS to modulate *FT* expression. *PLoS Genet.* 10:e1004617. doi: 10.1371/journal.pgen.1004617

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: 15 December 2021
doi: 10.3389/fpls.2021.761059

Profiling of H3K4me3 and H3K27me3 and Their Roles in Gene Subfunctionalization in Allotetraploid Cotton

Aicen Zhang[1†], Yangyang Wei[2†], Yining Shi[1†], Xiaojuan Deng[3†], Jingjing Gao[1†], Yilong Feng[1], Dongyang Zheng[1], Xuejiao Cheng[1], Zhaoguo Li[2], Tao Wang[2], Kunbo Wang[4], Fang Liu[4], Renhai Peng[2]* and Wenli Zhang[1]**

[1] State Key Laboratory for Crop Genetics and Germplasm Enhancement, JCIC-MCP, CIC-MCP, Nanjing Agricultural University, Nanjing, China, [2] Biological and Food Engineering, Anyang Institute of Technology, Anyang, China, [3] College of Agronomy, Xinjiang Agricultural University, Ürümqi, China, [4] Zhengzhou Research Base, State Key Laboratory of Cotton Biology, Zhengzhou University/Institute of Cotton Research, Chinese Academy of Agricultural Sciences, Anyang, China

OPEN ACCESS

Edited by:
Jie Song,
Imperial College London,
United Kingdom

Reviewed by:
Kai Wang,
Nantong University, China
Daisuke Miki,
Shanghai Center for Plant Stress Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Wenli Zhang
wzhang25@njau.edu.cn
Fang Liu
liufcri@163.comb
Renhai Peng
aydxprh@163.com

[†] These authors have contributed equally to this work

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

Received: 19 August 2021
Accepted: 17 November 2021
Published: 15 December 2021

Citation:
Zhang A, Wei Y, Shi Y, Deng X, Gao J, Feng Y, Zheng D, Cheng X, Li Z, Wang T, Wang K, Liu F, Peng R and Zhang W (2021) Profiling of H3K4me3 and H3K27me3 and Their Roles in Gene Subfunctionalization in Allotetraploid Cotton. Front. Plant Sci. 12:761059. doi: 10.3389/fpls.2021.761059

Cotton is an excellent model for studying crop polyploidization and domestication. Chromatin profiling helps to reveal how histone modifications are involved in controlling differential gene expression between A and D subgenomes in allotetraploid cotton. However, the detailed profiling and functional characterization of broad H3K4me3 and H3K27me3 are still understudied in cotton. In this study, we conducted H3K4me3- and H3K27me3-related ChIP-seq followed by comprehensively characterizing their roles in regulating gene transcription in cotton. We found that H3K4me3 and H3K27me3 exhibited active and repressive roles in regulating the expression of genes between A and D subgenomes, respectively. More importantly, H3K4me3 exhibited enrichment level-, position-, and distance-related impacts on expression levels of related genes. Distinct GO term enrichment occurred between A/D-specific and homeologous genes with broad H3K4me3 enrichment in promoters and gene bodies, suggesting that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome. An anticorrelation between H3K27me3 enrichment and expression levels of homeologous genes was more pronounced in the A subgenome relative to the D subgenome, reflecting distinct enrichment of H3K27me3 in homeologous genes between A and D subgenome. In addition, H3K4me3 and H3K27me3 marks can indirectly influence gene expression through regulatory networks with TF mediation. Thus, our study provides detailed insights into functions of H3K4me3 and H3K27me3 in regulating differential gene expression and subfunctionalization of homeologous genes, therefore serving as a driving force for polyploidization and domestication in cotton.

Keywords: H3K4me3, H3K27me3, ChIP-seq, gene expression, subfunctionalization, regulatory network, cotton

INTRODUCTION

The core nucleosome is the basic unit of chromatin. It comprises a core histone octamer, consisting of two copies of each histone H3, H4, H2A, and H2B, wrapped by approximately 147 bp of double-helix DNA (Kornberg, 1974). The exposed N-terminal tails of core histones, especially for histone H3, are subjected to various covalent modifications like methylation and acetylation with distinct

biological consequences (Fuchs et al., 2006; Kouzarides, 2007; Bannister and Kouzarides, 2011). In plants, individual histone modifications, especially combinatorial actions of multiple modifications, play essential roles during entire processes of normal growth and development (Berr et al., 2011; Deal and Henikoff, 2011; Boycheva et al., 2014; Han et al., 2019) and in response to diverse environmental cues (Kim J. M. et al., 2015; Ramirez-Prado et al., 2018; Alonso et al., 2019; Ueda and Seki, 2020). In particular, aberrant histone modifications frequently cause severe developmental defects in plants (Tian and Chen, 2001; Sanders et al., 2017; Deevy and Bracken, 2019), indicating histone modifications are indispensable for normal plant growth and development.

Histone methylation exhibits lysine site- and methylation degree-dependent effects on chromatin structure and gene transcription. Selective lysine residues can be mono-, di-, and trimethylated with distinct biological outcomes (Sims et al., 2003; Tariq and Paszkowski, 2004; Zhang et al., 2009; Liu et al., 2010; Stillman, 2018). H3K4me3 is a major active histone mark in eukaryotes (Santos-Rosa et al., 2002; Zhang et al., 2015a). By contrast, H3K27me3 is a repressive mark with significant biological relevance in eukaryotes, including plants (Zheng and Chen, 2011; Wiles and Selker, 2017).

So far, H3K4me3 and H3K27me3 are well-studied marks in plants (Deal and Henikoff, 2011). They play crucial roles in regulating gene expression during normal growth and development and responses to environmental cues in plants (Probst and Mittelsten Scheid, 2015; Chen D. H. et al., 2018). In general, H3K4me3 primarily distributes around the TSS or gene bodies of expressed genes, which is directly related to gene activation (Wang et al., 2009; Zhang et al., 2009; He et al., 2010). In contrast, H3K27me3 is primarily enriched in the promoters and gene bodies of repressed genes responsible for silencing conditionally expressed genes (Zhang et al., 2007; Wang et al., 2009; Deal and Henikoff, 2010; He et al., 2010). Moreover, the H3K27me3 mark has been recently found to be more associated with homeologs less expressed in polyploidy wheat (Ramirez-Gonzalez et al., 2018), indicating involvement of this mark in the expression bias of homeologs in the polyploidy genome.

Cotton (*Gossypium* spp.) is one of the major suppliers of natural textile fibers and oilseeds around the world. Allotetraploid cotton is an excellent model system for studying crop polyploidization and domestication. The availability of high-quality genome sequences for cultivated allotetraploid species and their wild relatives (Wang et al., 2012, 2019; Li F. et al., 2015; Zhang et al., 2015b; Chen et al., 2020) has promoted evolutionary and functional genomics studies (Zaidi et al., 2018), therefore, deepening our understanding of cotton biology and benefiting yield and fiber improvement in cotton breeding. However, when compared with tremendous progress in other model plants, like *Arabidopsis*, rice, and maize, histone modification-related epigenomic studies are largely understudied in cotton (Wang et al., 2016, 2017; Zheng et al., 2016; You et al., 2017; Tao et al., 2020), especially H3K27me3, a repressive mark for regulating genes involved in development and stress responses.

Here, we conducted H3K4me3 and H3K27me3 ChIP-seq assays using leaf tissue from the allotetraploid cotton cultivar *Gossypium hirsutum* (TM-1). We examined subgenomic distribution of both marks. In particular, we comprehensively investigated the roles of H3K4me3 and H3K27me3 marks in regulating the differential expression of genes and homeologous genes between the A and D subgenomes.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The allotetraploid cotton cultivar *G. hirsutum* was used in this study. Cotton seeds were soaked in water for 24 h, then transferred to soil in pots and continued to grow in a greenhouse with 60% humidity at 28°C/25°C and 16/8-h light/dark cycle. Two and three true leaves collected from 20-day-old plants were ground into a fine powder using liquid nitrogen. The ground powder can be used immediately or kept at −80°C for later use.

RNA-Seq

Total RNA was extracted from ground powder using TRIzol (Thermo Fisher Scientific). Total RNA was treated with DNaseI to completely remove contamination of genomic DNA. mRNA was enriched from DNaseI-treated RNA for library preparation and sequenced on Illumina HiSeq4000.

ChIP-Seq Assay

Twenty-day-old cotton leaves were used for nuclei preparation. The nuclei were extracted using a nuclei isolation buffer (NIB, pH 5.0, 1.0 M glucose, 0.1 M citric acid, 80 mM KCl, 10 mM EDTA, 1% Triton X-100 prepared fresh just before use). The nuclei were purified using 1 × nuclei washing buffer (NWB, pH 5.0, 1 M glucose, 0.1 M Na-citrate, 1%Triton X-100 prepared fresh just before use). The purified nuclei were resuspended using 600 μl MNB buffer (50%, w/v sucrose, 50 mM Tris–HCl, pH 7.5, 4 mM $MgCl_2$, 1 mM $CaCl_2$) for MNase digestion at 37°C for 10 min. A digestion mix was pelleted at 13,000 rpm for 10 min at 4°C and the supernatant was transferred into a new 1.5 ml tube. The digested nuclei pellet was resuspended using lysis buffer (1 mM Tris–HCl, pH 7.5, 0.1 mM PMSF, 2%, v/v Complete Mini) and left on ice for 1 h. After centrifugation, the supernatant was transferred into the 1.5 ml tube containing digested chromatin. ChIP incubation buffer was added to the digested chromatin to make a total of 1.7 ml. The remaining steps were conducted following the published procedures (Zhang et al., 2012), namely, antibody incubation followed by adding protein A-sepharose beads, bead washing, elution, and purification of ChIPed DNA for library preparation. The prepared libraries were finally sequenced on the Illumina platform (Illumina HiSeq4000).

Processing of Sequencing Data

Raw reads of all sequencing data were trimmed using fastp (Chen S. et al., 2018) based on the quality value ($Q \geq 25$) and read length (≥20 bp). The trimmed reads from RNA-seq and ChIP-seq were mapped to the *G. hirsutum* reference genome (Zhang et al., 2015b) with Hisat2 (Kim D. et al., 2015) and Bowtie2 aligner (Langmead and Salzberg, 2012), respectively. The remaining RNA-seq data of different tissues were obtained from

the previously published data (Zhang et al., 2015b), and processed according to the procedures described above. For ChIP-seq data, any PCR duplicates were removed using Picard. Aligned reads with mapping quality (MapQ) less than 30 were removed using samtools (Li et al., 2009). The bam files were converted to bigwig file and normalized by RPKM (Reads Per Kilobase per Million mapped reads) using deeptools (Ramirez et al., 2016), then Integrative Genomics Viewer (IGV) (Thorvaldsdottir et al., 2013) was used to visualize read distribution across the genome.

Identification of Differentially Expressed Homeologous Genes

Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) computed from StringTie (Pertea et al., 2015) was used to measure the expression level of each gene. Gene annotations were obtained from CottonGen (Yu et al., 2014). Only genes with FPKM $\geq$ 1 were considered as expressed ones. Homeologous gene pairs with FPKM $\geq$ 1 in at least one of the subgenomes were used for further analyses. The homeologous gene pairs (**Supplementary Table 1**) were identified using reciprocal BLAST hits between A and D homeologs as previously reported (Zhang et al., 2015b). Differentially expressed homeologous genes were analyzed using the limma R package. The *P* values were adjusted using the BH method at $\alpha = 0.05$ (Benjamini and Hochberg, 1995). Corrected *P* values of 0.05 and $\log_2$ (fold-change) values of 1 were set as the threshold for assessing differential expression levels.

ChIP-Seq Data Analyses

In this study, we only used rep I data of H3K4me3 and H3K27me3 for peak calling and subsequent analyses, and we used rep II data to validate the reproducibility of peaks for each mark. H3K27me3 peaks were called with "--broad" parameter on (-f BAM -g 2.3e+9 --nomodel -q 0.01 --broad --broad-cutoff 0.1) and H3K4me3 peaks were called with off (-f BAM -g 2.3e+9 --nomodel -q 0.01) using the MACS2 software (Zhang et al., 2008). The ChIPpeakAnno package (Zhu et al., 2010) was used for peak annotation. Bedtools (Quinlan and Hall, 2010) were used to correlate peaks with genome loci, genes (including 1,000 bp region upstream of the TSS and 1,000 bp region downstream of the TTS) overlapping H3K4me3 or H3K27me3 peaks were considered as peak-related genes. A custom script was used to calculate normalized read counts of related genes.

Regulatory Network Analyses

The biased genes between the A and D subgenomes were chosen to construct gene co-expression networks using the R package WGCNA (Langfelder and Horvath, 2008). The blockwiseModules function was used for network construction with the following parameters: power = 16, minModuleSize = 30, mergeCutHeight = 0.25, corType = "pearson." The same expression matrix used for WGCNA analyses was used for regulatory network analyses with the R package GENIE3 (Huynh-Thu et al., 2010). All TF annotations were obtained from PlnTFDB version 3.0 (Jin et al., 2014) and CottonFGD[1].

[1]https://cottonfgd.org/

Gene Ontology Enrichment Analyses

Gene Ontology enrichment analyses were conducted using agriGO v2.0 (Tian et al., 2017). GO terms with an FDR less than 0.05 were considered as being significantly enriched.

Bimolecular Fluorescent Complimentary (BiFC) Assay

To verify protein–protein interactions among PRE6, Rf2b, and bHLH63 as shown in **Figure 4B**, we performed transient transformation related Bimolecular Fluorescent Complimentary (BiFC) assays using *Nicotiana benthamiana*. To generate vectors of pXY104 with the expression cassette of 35S: *PRE6-cYFP*, R*f2b-cYFP* and *bHLH63-cYFP*, we fused the coding sequences of PRE6, Rf2b, and bHLH63 to the C-terminal half of YFP. Meanwhile, we tagged each gene with the N-terminal half of YFP to generate vectors of pXY104 with the expression cassette of 35S:*PRE6-nYFP*, *Rf2b-nYFP* and *bHLH63-nYFP* (**Supplementary Table 2**). All plasmids were individually transferred into *Agrobacterium tumefactions* (GV3101), and cultured in LB medium at 28°C until OD600 reached 1.2–1.5. After collecting the strains by centrifugation at 5,000 rpm at 4°C, the strains were individually suspended in injection buffer (10 mM MES-KOH, 10 mM $MgCl_2$, 40 μM AS, pH = 5.7). After incubation at room temperature for 2 h, the bacteria containing cYFP/nYFP, Rf2b-cYFP/Rf2b-nYFP, Rf2b-cYFP/PRE6-nYFP, Rf2b-cYFP/bHLH63-nYFP, PRE6-cYFP/PRE6-nYFP, PRE6-cYFP/bHLH63-nYFP, and bHLH63-cYFP/bHLH63-nYFP were combined in 1:1 ratio, and were then individually co-injected into young and healthy *N. benthamiana* leaves for transient expression. The fluorescent signal was monitored and recorded using the Leica DMi8 confocal laser scanning microscope at 72 h post-infiltration, at least three images per injection. The functional validation for each combination was repeated by at least three separate injections.

RESULTS

Global Distribution of H3K4me3 and H3K27me3 in Cotton

To examine global distributions of H3K4me3 and H3K27me3 in the allotetraploid cotton, we conducted each mark-related ChIP-seq (**Supplementary Table 3**). We found that biological replicates of each mark were well correlated (**Supplementary Figure 1**). After data processing and peak calling, we identified total 75,516 H3K4me3 peaks (37,499 peaks in A subgenome and 38,017 peaks in D subgenome), 22,543 H3K27me3 peaks (10,322 peaks in A subgenome and 12,221 peaks in D subgenome) and 10,641 loci with at least 1bp overlap between H3K4me3 and H3K27me3 peaks (4,908 loci in A subgenome and 5,733 loci in D subgenome) (**Figures 1A,B**). The representative IGV snapshots (**Figure 1B**) show reproducible distribution of each mark in the A and D subgenomes. We then plotted normalized H3K4me3 and H3K27me3 ChIP-seq read counts across ± 1kb of the TSS and the TTS of all genes in the A and D subgenomes. We observed a distinct genic distribution of each mark, H3K4me3 was highly

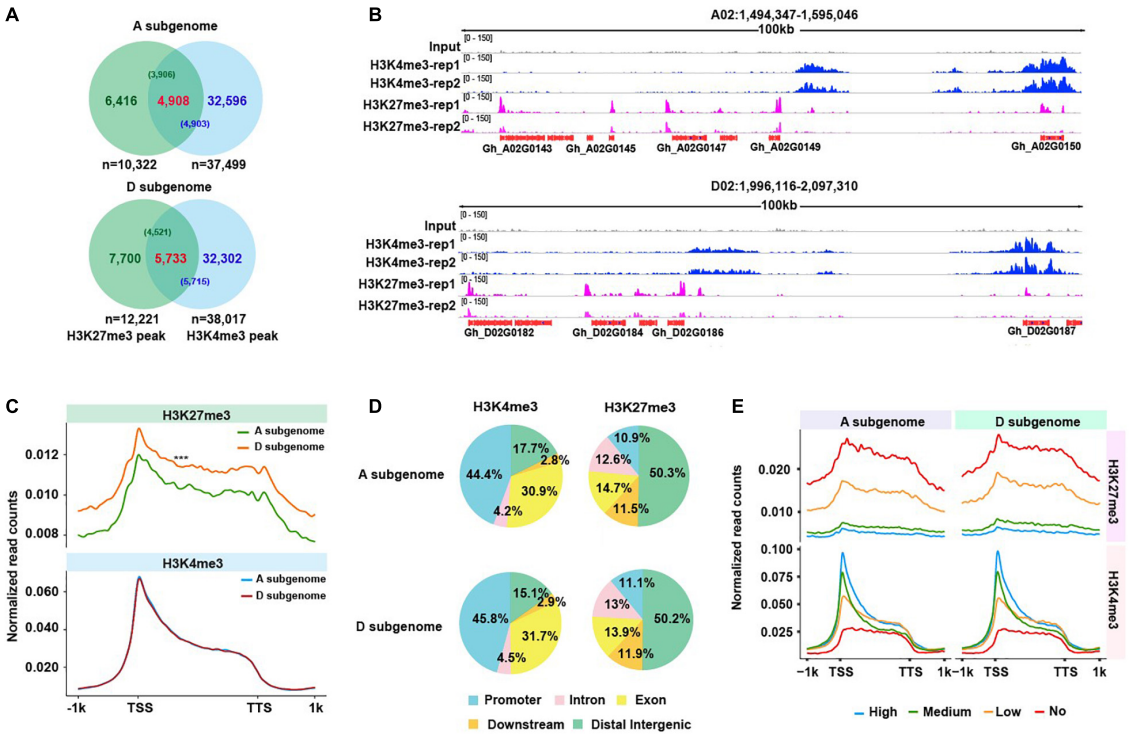

FIGURE 1 | Distribution of H3K4me3 and H3K27me3 marks. **(A)** Venn plots illustrating overlaps of H3K4me3 and H3K27me3 peaks, and H3K4me3/H3K27me3 bivalents identified from A and D subgenome, respectively. The number in each bracket represents the number of H3K4me3 (or H3K27me3) peaks that overlap with H3K27me3 (or H3K4me3). **(B)** Representative IGV snapshots across a 100 kb window from chromosome A02 and chromosome D02 show the enrichment of H3K4me3 and H3K27me3 peaks in the A and D subgenomes. **(C)** Curve plots show the profile of normalized H3K4me3 and H3K27me3 read counts from 1 kb upstream of the TSSs to 1 kb downstream of the TTSs across all genes from two subgenomes. A significance test was determined using the Wilcoxon rank-sum test, $^{***}p < 0.001$. **(D)** Distribution of H3K4me3 and H3K27me3 in different functional sub-genomic annotations, namely, promoters (upstream 2 kb), exons, introns, downstream (2 kb) and distal intergenic regions. **(E)** Curve plots show the profile of normalized H3K4me3 and H3K27me3 read counts from 1 kb upstream of the TSS to 1 kb downstream of the TTS of all genes from the A and D subgenomes. All genes were divided into four subtypes according to their FPKM values (high, medium, low, and no expression).

enriched immediately downstream of the TSS and extended to the whole gene body, by contrast, H3K27me3 primarily covered the whole gene body (**Figure 1C**). Strikingly, we found that distributions of normalized read counts for H3K4me3 were similar between A and D subgenome, whereas distributions of normalized read counts for H3K27me3 were higher in D subgenome relative to A subgenome.

To visualize the distribution of each mark in the A and D subgenomes, we partitioned each subgenome into five functionally annotated subregions, namely, promoters, exons, introns, downstream of TTS, and distal intergenic regions. We observed subtle differences in H3K4me3 but similar distributions for H3K27me3 between A and the D subgenomes (**Figure 1D**). Compared to the D subgenome, A subgenome had approximately 2% more H3K4me3 distributed in distal intergenic regions, and 1.4% less H3K4me3 distributed in promoters. Moreover, a distinct subgenomic distribution was observed for each mark, H3K4me3 exhibited the highest distribution in promoters, while H3K27me3 had the highest distribution in distal intergenic regions, suggesting a potential mark-dependent functional divergence. To assess an association between H3K4me3 or H3K27me3 enrichment levels and gene expression levels in the A and D subgenome, we classified all genes in the A and D subgenome into four subtypes (high, medium, low, and no expression) according to FPKM values. We then plotted normalized H3K4me3 or H3K27me3 read counts around ± 1kb of the TSS and the TTS of genes. We found that H3K4me3 enrichment levels, indicative of normalized read counts, exhibited a positive correlation with gene expression levels, whereas H3K27me3 enrichment levels were anti-correlated with gene expression levels in A and D subgenome (**Figure 1E**). Consistent with previous reports in other plant species, H3K4me3 and H3K27me3 also have contrasting roles in regulating gene expression in cotton, the former can facilitate gene expression whereas the latter usually suppresses gene expression.

Distinct Roles of H3K4me3 and H3K27me3 in Regulating Expression of Homeologous Genes

After comparing expression levels of genes between the A and D subgenomes, we observed that genes in A subgenome generally expressed more than those in D subgenome (**Figure 2A**),

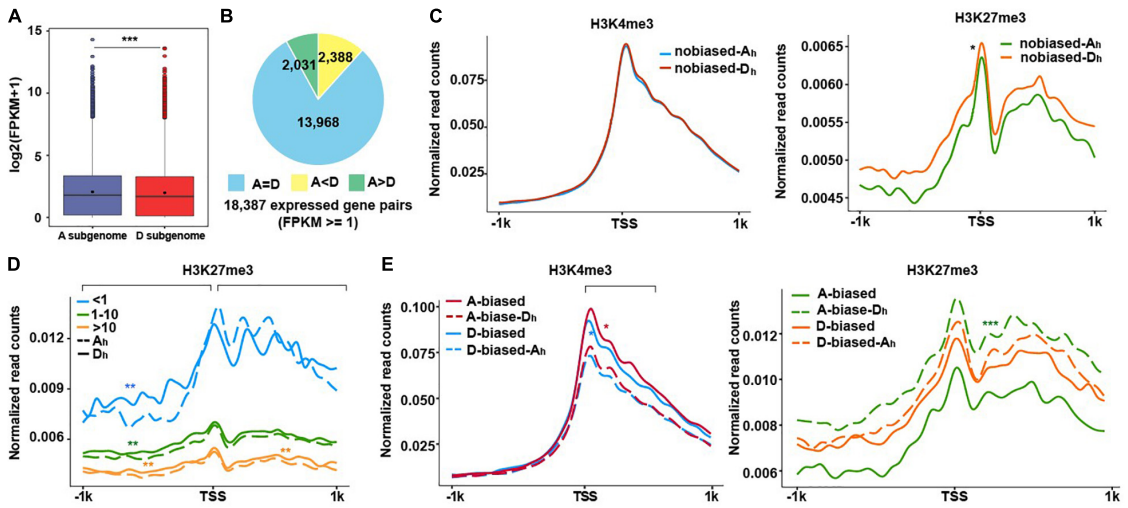

FIGURE 2 | Subgenomic variations in relationships between H3K4me3/H3K27me3 and gene expression levels. **(A)** The boxplot show the average gene expression levels from A and D subgenome. A significance test was determined using the Wilcoxon rank-sum test, $^{***}p < 0.001$. **(B)** Pie charts show the number of homeologs with equal (A = D), A homeolog-biased (A > D), and D homeolog-biased (A < D) expression. **(C–E)** The profiles of normalized H3K4me3 and H3K27me3 read counts from ± 1 kb around TSS of the homeologous genes from two subgenomes. H3K4me3 (left) and H3K27me3 (right) normalized read counts for no biased gene pairs **(C)**. H3K27me3 normalized read counts for no biased genes that were divided into three groups (FPKM < 1, 1–10, and ≥10) **(D)**. H3K4me3 (left) and H3K27me3 (right) normalized read counts for biased gene pairs **(E)**. The significance test was determined using the Wilcoxon rank-sum test, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

indicating subgenome-biased expression in cotton. To examine how H3K4me3 and H3K27me3 are involved in regulating the differential expression of homeologous genes, we classified 18,387 expressed homeologous gene pairs into three subtypes: non-biased expression, A = D (n = 13,968), representing equal expression levels between A and D; A-biased expression, A > D (n = 2,031), representing genes expressed more in A homeologs and D-biased expression, D > A (n = 2,388), representing genes expressed more in D homeologs (**Figure 2B**; **Supplementary Figure 2**; **Supplementary Table 4**). After comparing normalized read counts, we found that there was no difference for H3K4me3 for no biased A_h and D_h (A_h and D_h represent A homeologs and D homeologs, respectively), whereas more H3K27me3 occurred in no biased-D_h than no biased-A_h, the counterparts in the A subgenome (**Figure 2C**).

To assess if read density changes of H3K27me3 in no biased homeologous genes are possibly related to gene expression levels, we classified no biased homeologous genes into three subtypes: FPKM < 1, 1–10, and > 10 and conducted similar plotting assays as **Figure 2C**. When compared with genes in the D subgenome, we observed that non-expressed genes (FPKM < 1) in A exhibited less H3K27me3 reads distributed at the upstream of TSSs, but more H3K27me3 reads distributed at the downstream of TSSs. For expressed genes (FPKM ≥ 1), H3K27me3 reads tended to have more in the D subgenome than the A subgenome across all regions examined (**Figure 2D**).

We conducted similar plotting analyses for A- and D-biased expressed genes. As shown in **Figure 2E** (left), H3K4me3 exhibited higher enrichment in A-biased and D-biased genes compared to their respective counterpart of homeologous genes (A-biased-D_h and D-biased-A_h), indicating that H3K4me3 enrichment levels are directly correlated with gene expression levels. By contrast, H3K27me3 was less enriched in A-biased and D-biased genes compared to the corresponding A-biased-D_h and D-biased-A_h, respectively (**Figure 2E** right), exhibiting an anticorrelation between H3K27me3 enrichment levels and gene expression levels. After a careful examination, we found that the difference in H3K27me3 between A-biased and A-biased-D_h was more pronounced than that between D-biased and D-biased-A_h. These results showed that an anti-correlation between H3K27me3 enrichment levels and expression levels of homeologous genes was more pronounced in the A subgenome relative to the D subgenome, reflecting distinct enrichment of H3K27me3 in homeologous genes between A and D subgenome. Similarly, the roles of H3K27me3 in regulating biased expression of homeologs have been investigated in polyploidy wheat (Ramirez-Gonzalez et al., 2018), indicating potential common roles of H3K27me3 in regulating differential expression of homeologs in polyploidy plants.

Biological Implications of Genes With Broad H3K4me3

It has been reported that genes with broad H3K4me3 enrichment have some particular biological implications, such as determination of cell identity, regulation of expression of cell-type-specific tumor suppressors, regulation of expression of genes responsible for gamete development, and pre-implantation in mammalians (Benayoun et al., 2014; Chen et al., 2015; Dahl et al., 2016; Liu et al., 2016; Lv and Chen, 2016; Zhang et al., 2016) and potential roles in photosynthesis in *Arabidopsis* (Brusslan et al., 2015). To interrogate if genes marked with broad H3K4me3 enrichment have any distinct biological

implications in cotton, we conducted *k*-means clustering assay using H3K4me3 associated genes in the A and D subgenomes. We obtained five clusters of genes with distinct H3K4me3 ChIP-seq read distribution (**Figures 3A,B**). There were two types of genes (Cluster 1 and 3) with broad H3K4me3 enrichment in promoter and gene body regions, respectively. After comparing gene expression levels and gene length in each cluster, genes in each cluster between the A and D subgenomes had overall similar mean expression levels (**Figure 3C**) and mean gene length (**Supplementary Figure 3A**). We found that genes in Cluster 4 had the highest mean expression levels and gene length, whereas genes in Cluster 5 had the highest gene length but the lowest mean expression levels in both subgenomes and genes in Cluster 3 had the shortest gene length (**Figure 3C**; **Supplementary Figure 3A** and **Supplementary Table 5**). These results suggest that the impacts of H3K4me3 on gene expression may depend on its enrichment levels, position, and distance from H3K4me3 to the TSS.

We then conducted GO term enrichment analyses using genes in Clusters 1 and 3 in A and D subgenome. For the genes in Cluster 1, we found that majority of GO terms were common between A and D subgenome despite several distinct GO terms occurred between A and D subgenome. For instance, the genes in D subgenome were more enriched in macromolecular complex/metabolic processes, cellular processes and transducer activities while the genes in A subgenome were more enriched in protein binding (**Supplementary Figure 4**). We further divided the genes in Cluster 1 into A or D subgenome-specific and homeologous gene pairs (**Supplementary Figure 3B**) for re-conducting GO term enrichment assays. We observed distinct GO terms occurred between A- and D-specific genes (**Figure 3D**). A subgenome-only genes were more enriched in cellular component category but less enriched in molecular function and biological process categories compared to D subgenome-only genes. Homeologous gene pairs were more involved in the cellular component category but less in the molecular function category.

As illustrated in **Figure 3A** (heatmap) and **Supplementary Figure 3A** (boxplot), the genes in Cluster 3 contain broad H3K4me3 mark covering downstream 2 kb of the TSSs, but have the shortest gene length. To specifically look into biological relevance of genes with broad H3K4me3 enriched in gene body instead of extending to downstream of the genes with length less than 2 kb, we conducted similar GO term analyses using the genes in Cluster 3 with length greater than 2 kb in A and D subgenome (**Supplementary Figures 5A,B**). We observed subtle differences in GO terms between A and D genome. Compared to the genes in A subgenome, the genes in D subgenome were more enriched in signal transducer activity, carbohydrate binding, cell communication and pollination and pollen-pistil interaction (**Supplementary Figure 5C**). After dividing the genes specific for A or D subgenome and

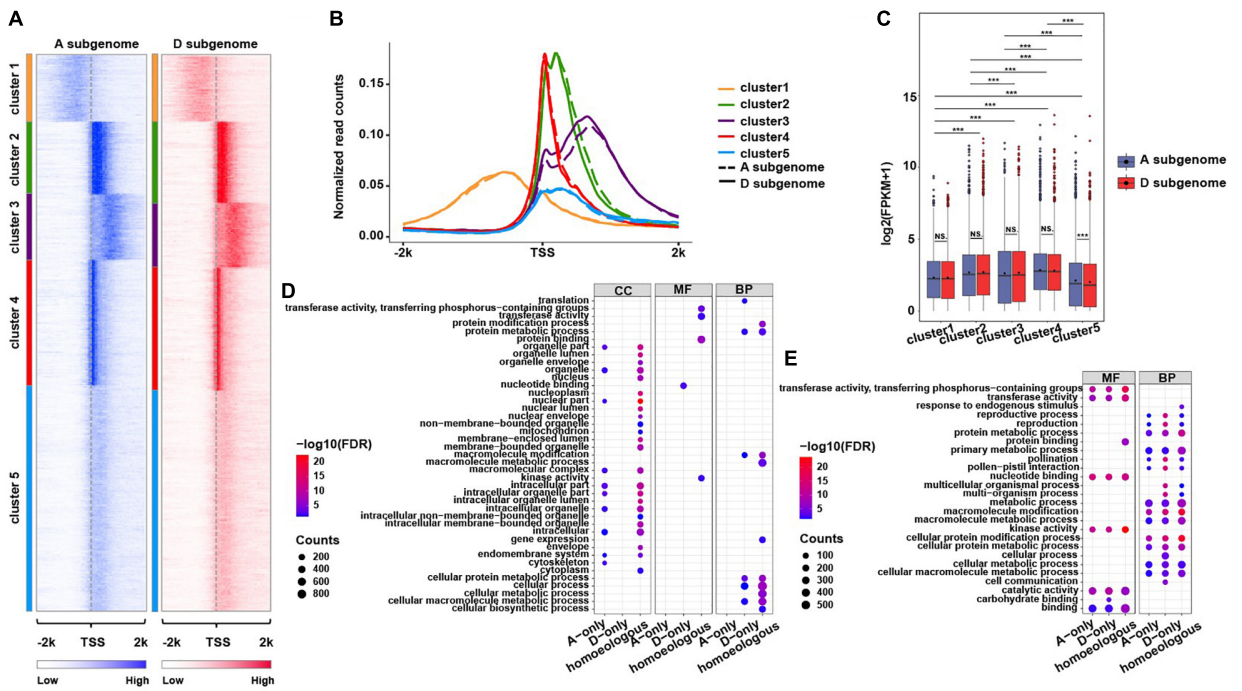

FIGURE 3 | Distribution of H3K4me3 mark around TSSs of the genes in two subgenomes. **(A)** Heatmaps showing five *k*-means clusters for H3K4me3 distributed around ± 2 kb of TSSs of the overlapping genes in A and D subgenome. The color represents H3K4me3 enrichment levels. **(B)** The profiles of normalized H3K4me3 read counts around ± 2 kb of TSSs of the genes from each cluster characterized in **(A)**. **(C)** The boxplots show expression levels of the genes from each cluster in two subgenomes. A significance test was determined using the Wilcoxon rank-sum test, $^{***}p < 0.001$. **(D,E)** Functional GO term enrichment analyses of the genes from Cluster1 **(D)** and Cluster3 **(E)**, the related genes were divided into three types, A subgenome-only, homeologous (homeologous gene pairs in the same cluster in A and D subgenome) and D subgenome-only, the size of each dot represents the number of genes, and the color key indicates – log10 (FDR).

homeologous gene pairs (**Supplementary Figure 5D**), we found the genes in D subgenome only were more enriched in functions associated with carbohydrate binding, metabolic processes, pollination, and pollen–pistil interaction as compared to the genes in A subgenome only (**Figure 3E**). To test if genes with broad H3K4me3 mark exhibit tissue-specific differential expression, we conducted *k*-means clustering analyses using RNA-seq data derived from 12 distinct tissues (the public data), we found that the genes with broad H3K4me3 enrichment exhibited tissue-specific expression profiles, including significantly highly expressed genes in stamen and petal (**Supplementary Figures 6A,B**), suggesting that broad H3K4me3-marked genes may function in reproductive stage for flower development. These analyses suggest that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome.

Involvement of H3K4me3 and H3K27me3 in Regulating Gene Expression Through TF-Mediated Regulatory Network

After specifically examining TFs with subgenome-related differential expression, we detected 180 and 204 TFs with biased expression in A and D subgenome, respectively (**Supplementary Figure 7** and **Supplementary Table 6**). A- and D-biased TFs associated with H3K4me3-only, H3K27me3-only, and H3K4me3/H3K37me3 mark were summarized in **Supplementary Tables 6, 7**. To interrogate if H3K4me3 and H3K27me3 function in regulating gene expression through TF-mediated regulatory networks, we conducted a co-expression assay with the WGCNA R package, an expression matrix of biased genes across 12 tissues was used for the downstream analyses. We obtained 21 co-expression modules (**Supplementary Figure 8A**). Genes in the blue module exhibited a high association in the leaf tissue (**Supplementary Figure 8A**) and eigengenes in the blue module were specifically expressed in the leaf tissue (**Supplementary Figure 8B**).

Hub genes have been reported to be essential for maintaining the structure of the corresponding module and network (Li Y. et al., 2015; van Dam et al., 2018). We found that the blue module contained 22 hub TFs with a module membership value (|kME|) > 0.9, which is designed as the Pearson's correlation coefficient between the expression of a gene and a given module epigengene (Langfelder and Horvath, 2008). The 22 hub TFs contained 3 TFs enriched with H3K27me3, 4 TFs associated with H3K4me3/H3K27me3 marks, and the rest (15 TFs) enriched with H3K4me3 (**Supplementary Figure 9** and **Supplementary Table 8**).

To further infer interactions between hub TFs, we built a gene regulatory network with 22 hub TFs as candidate regulators

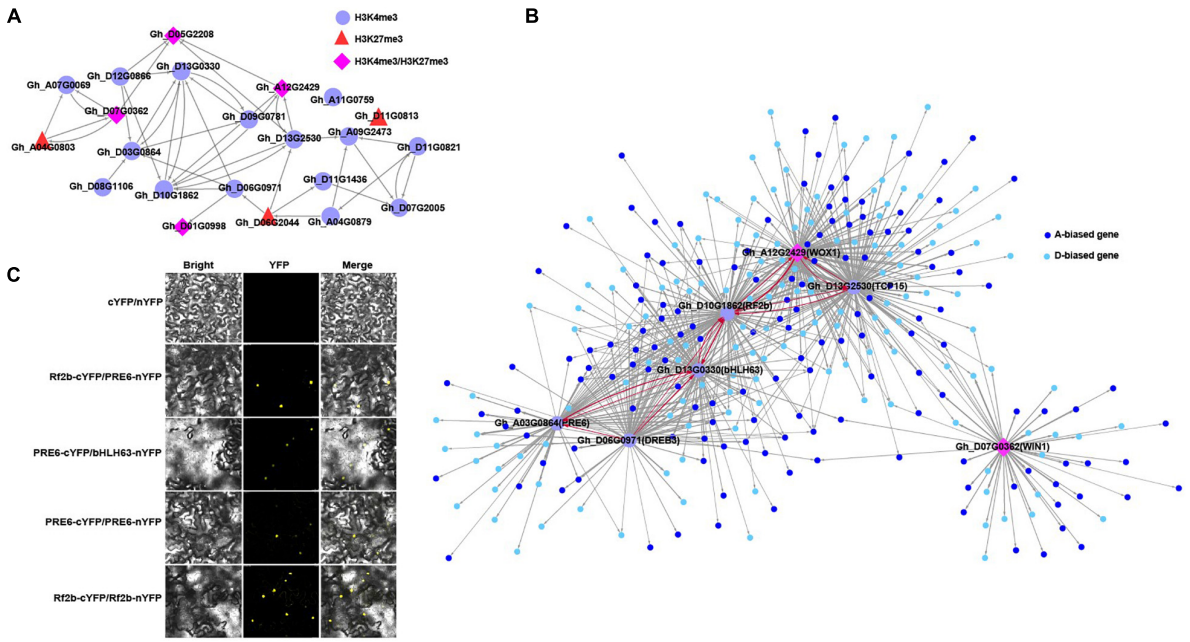

FIGURE 4 | Regulatory network in cotton leaf tissue. **(A)** Overview of predicted regulatory network containing 22 hub TFs with different marks in the blue module. The TFs with H3K4me3, H3K27me3, and H3K4me3/H3K27me3 marks were indicated using different shapes and colors; the arrow represents the direction of regulation. **(B)** The directional regulatory network show the top 7 TFs with the most edges in **(A)** and their predicted regulated genes (small nodes with different colors), the arrow represents the direction of regulation, the red line indicates interactions between TFs, thickness of the line represents the weight value. **(C)** BiFC analyses of protein–protein interactions of PRE6, Rf2b and bHLH63. Bright, YFP, and Merge (bright-field and yellow fluorescent), Rf2b-cYFP/PRE6-nYFP: the C-terminal half of YFP was fused to the C-terminal of PRE6, Rf2b and bHLH63 to generate PRE6-cYFP, Rf2b-cYFP, and Bhlh63-cYFP, whereas the N-terminal half of YFP was fused to the N-terminal of PRE6, Rf2b, and bHLH63 to generate PRE6-nYFP, Rf2b-cYFP, and bHLH63-cYFP; *Agrobacterium* combination of Rf2b-cYFP/PRE6-nYFP, PRE6-cYFP/bHLH63-nYFP, PRE6-cYFP/PRE6-nYFP, and Rf2b-cYFP/Rf2b-nYFP were individually co-injected into *N. benthamiana* leaves for transient expression. The fluorescent signal was monitored and recorded by confocal microscopy at 72 h post-infiltration. Bar: 25 μm.

for the expression of other genes using the GENIE3 R package (Huynh-Thu et al., 2010; **Figure 4A**). To clearly show predicted regulatory relationships between TFs, the edges among other genes involved in the network were not displayed. Subsequently, the top 7 TFs with the most edges as displayed in **Figure 4A** and their regulated genes preferentially expressed in either the A or D subgenome were illustrated in **Figure 4B**. Strong interactions between TFs occurred among PRE6, BHLH63, RF2b, WOX1, and TCP15. PRE6, BHLH63, and RF2b were regulated by DREB3. To validate the accuracy of the network, we conducted a BiFC assay for Rf2b, bHLH63, and PEE6. Protein interactions were detected between Rf2b/bHLH63 and PEE6, and self-interaction was observed for PRE6 and Rf2b proteins (**Figure 4C**).

Functions of each TF ortholog have been documented in other plants such as *Arabidopsis* and rice. For example, it has been documented that TCP15 acts as a repressor of auxin biosynthesis and functions in the regulation of *Arabidopsis* gynoecium development (Lucero et al., 2015). DREB3 is a transcriptional activator functioning in abiotic stress responses in plants (Niu et al., 2020). bHLH63 also known as *CRYPTOCHROME-IN-TERACTING* bHLH1 and bHLH100, is mainly involved in embryo suspensor and postembryonic development in *Arabidopsis* (Liu et al., 2008; Radoeva et al., 2019).

Collectively, our analyses indicate that, in addition, to directly affecting the expression of overlapping genes, H3K4me3, H3K27me3, and H3K4me3/H3K37me3 marks can indirectly influence gene expression through TF-mediated regulatory networks in the leaf tissue.

DISCUSSION

Similar to previous findings in plants (Zhang et al., 2007, 2009; Wang et al., 2009; He et al., 2010; Deal and Henikoff, 2011), our study indicated that H3K4me3 and H3K27me3 are highly enriched in gene bodies, H3K4me3 is an active mark that directly correlates with expression levels of genes, whereas H3K27me3 is a repressive mark that anti-correlates with expression levels of genes in allotetraploid cotton. It has been documented that H3K4me3 is involved in the biased expression of homeologous genes in allotetraploid cotton root (Zheng et al., 2016), and differential enrichment of H3K4me3 is responsible for transcriptional changes of genes associated with cotton development and evolution (You et al., 2017). However, our study for the first time characterized possible roles of H3K27me3, and broad H3K4me3 in differentially regulating gene expression between A and D subgenome in cotton.

Roles of H3K27me3 and Broad H3K4me3 Enrichment in Gene Transcription in Cotton

Compared with extensive H3K27me3 studies in plant development and stress responses in other plant species (Zheng and Chen, 2011; Gan et al., 2015; Chang et al., 2020), the roles of H3K27me3 in cotton are still much less studied. In addition to the overall repressive role of H3K27me3 in regulating gene expression in cotton, our study showed that distinct enrichment of H3K27me3 in homeologous genes occurred between the A and D subgenomes (**Figures 2C–E**), since H3K27me3 enrichment in the A subgenome displayed a more pronounced anti-correlation with expression levels of homeologous genes as compared to the D subgenome.

Biological functions of broad H3K4me3 enrichment have been well studied in mammals, namely, cell identity (Benayoun et al., 2014), transcription of tumor suppressor genes (Chen et al., 2015; Dhar et al., 2018), transcription of genes in pre-implantation development, and embryonic stem cell differentiation (Liu et al., 2016), however, it is poorly understood in plants. In addition to the active roles of H3K4me3 in regulating gene expression, our study showed that the effects of H3K4me3 on expression levels of overlapping genes were related to their enrichment levels, relative position, and distance around the TSS in cotton (**Figures 3A–C**). Importantly, we found that A- and D-specific genes and homeologous genes with broad H3K4me3 in promoters and gene bodies were potentially involved in differential biological relevance and tissue-specific expression, suggesting that broad H3K4me3-marked genes might have some unique biological functions between A and D subgenome. For example, A-specific genes with broad H3K4me3 in promoters had more enriched GO terms relative to those D-specific genes, moreover, they had distinct GO terms associated with molecular functions and biological processes as compared to the corresponding genes with broad H3K4me3 enrichment in gene bodies (**Figures 3D,E**). It has been reported that genes with broad H3K4me3 enrichment have enriched GO terms associated with photosynthesis in *Arabidopsis* (Brusslan et al., 2015). Thus, our study provides further insights into the roles of broad H3K4me3 enrichment in subfunctionalization of homeologous genes in cotton.

Impacts of H3K4me3, H3K27me3 and H3K4me3/H3K27me3 Marks on Gene Transcription Through the Regulatory Network

In addition to direct impacts of H3K4me3 and H3K27me3 on transcription of overlapping genes in two subgenomes of cotton, our directional regulatory network related to TF indicated that both marks acted individually or in combined actions to indirectly regulate expression of co-expressed A- or D-biased genes through interacting between H3K4me3 or H3K27me3 mark overlapping TFs. For instance, predicted mutual interactions occurred between BHLH63 and PRE6 or RF2b, and between RF2b and TCP15, some of which were validated using a bimolecular fluorescence complementation assay. TCP 14 and TCP15 can regulate internode length and leaf shape in *Arabidopsis* through modulating cell proliferation (Kieffer et al., 2011). WOX1 is a key regulator during meristem development in *Arabidopsis* (Zhang et al., 2011). RF2b is a bZIP protein functioning in symptom development of rice tungro disease through interacting with RF2a (Dai et al., 2004). Functions of RF2b in leaf and root development in *Arabidopsis* are mediated by bZIP29 (Van Leene et al., 2016). PRE6 is a

paclobutrazol resistance protein belonging to non-DNA binding basic helix–loop–helix transcription factor, which has been reported to be involved in phytohormone signaling, such as GA, auxin, and BR and light responses in *Arabidopsis* (Gommers et al., 2017; Zheng et al., 2017). Gene regulatory network has already been applied to predict key nitrogen regulators, followed by successful experimental validation in rice (Ueda et al., 2020). We further extended the gene regulatory network to infer direct or indirect interactions between genes or TFs, which are responsible for differential expression of subgenome specific genes in allotetraploid cotton.

Collectively, our study provides evidence to indicate direct and indirect impacts of H3K4me3 and H3K27me3 on differential transcription of genes or homeologous gene pairs between A and D subgenome in cotton leaf tissue. In particular, the involvement of typically repressive mark H3K27me3 in expression bias of the homoeologs in cotton is still understudied. Therefore, we provide further evidence showing the involvement of H3K27me3 individually or in combination with H3K4me3 in regulating differential expression of homeologous gene pairs between the A and D subgenomes. Thus, H3K27me3 individually or coordinated with H3K4me3 could play important roles in genomic evolution and/or domestication through controlling bias expression of the homoeologs with some specific biological relevance in allotetraploid cotton. It has been reported that methylated genes evolve faster than unmethylated genes, and changes in DNA methylation and H3K4me3 enrichment are directly associated with expression bias of the homoeologs in allotetraploid cotton and their relatives (Song et al., 2017), thereby both epigenetic marks possibly functioning in gene domestication, including flowering time-related gene *GhCOL2* in cotton. Profiling of H3K4me3 and H3K27me3 marks in ancestors and relatives of allotetraploid cotton could help to address epigenetic regulatory roles underlying polyploidization and domestication in cotton.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author (WZ, wzhang25@njau.edu.cn), upon request. The raw sequencing data are deposited in the NCBI Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE165245.

AUTHOR CONTRIBUTIONS

WZ conceived and designed the study and wrote the manuscript with contributions from all the authors. YS, JG, YW, and XD conducted the experiments. AZ analyzed the data. YF, TW, and XC helped with plant growth. RP and ZL helped with the generation of RNA-seq data. FL and KW supervised the experiments. YS, JG, YW, AZ, and WZ interpreted the results. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (32070561, U20A2030, and 31471548), Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX21_0602), the Fundamental Research Funds for the Central Universities (KYYJ 201808), the State Key Laboratory of Cotton Biology Open Fund (CB2019A06), the Postdoctoral Fund of Anyang Institute of Technology (BHJ2019009), and the Innovation Practice Base Postdoctoral Fund of Henan Province (2019, to YW).

ACKNOWLEDGMENTS

The authors thank the Bioinformatics Center in Nanjing Agricultural University for providing facilities to assist sequencing data analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.761059/full#supplementary-material

Supplementary Table 1 | The homeologous gene pairs between A and D subgenome (in a separate Excel file).

Supplementary Table 2 | Primer information was used in this study.

Supplementary Table 3 | Summary of sequencing data.

Supplementary Table 4 | The expression levels and fold changes of differentially expressed homeologous gene pairs (in a separate Excel file).

Supplementary Table 5 | The gene length and expression levels of 5 broad H3K4me3 marked genes from five clusters (in a separate Excel file).

Supplementary Table 6 | Details of A- or D-biased TFs and 182 biased TFs with H3K4me3, H3K27me3, or H3K4me3/H3K27me3 marks (in a separate Excel file).

Supplementary Table 7 | Summary of A- and D-biased TFs associated with H3K4me3, H3K27me3, and both marks.

Supplementary Table 8 | Genes and their kME values in the leaf-related blue module (in a separate Excel file).

REFERENCES

Alonso, C., Ramos-Cruz, D., and Becker, C. (2019). The role of plant epigenetics in biotic interactions. *New Phytol.* 221, 731–737. doi: 10.1111/nph.15408

Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395. doi: 10.1038/cr.2011.22

Benayoun, B. A., Pollina, E. A., Ucar, D., Mahmoudi, S., Karra, K., Wong, E. D., et al. (2014). H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* 158, 673–688. doi: 10.1016/j.cell.2014.06.027

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc., Ser. B, Methodol.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Berr, A., Shafiq, S., and Shen, W. H. (2011). Histone modifications in transcriptional activation during plant development. *Biochim. Biophys. Acta* 1809, 567–576. doi: 10.1016/j.bbagrm.2011.07.001

Boycheva, I., Vassileva, V., and Iantcheva, A. (2014). Histone acetyltransferases in plant development and plasticity. *Curr. Genomics* 15, 28–37. doi: 10.2174/1389202915011403061127 42

Brusslan, J. A., Bonora, G., Rus-Canterbury, A. M., Tariq, F., Jaroszewicz, A., and Pellegrini, M. (2015). A Genome-wide chronological study of gene expression and two histone modifications, H3K4me3 and H3K9ac, during developmental leaf senescence. *Plant Physiol.* 168, 1246–1261. doi: 10.1104/pp.114.252999

Chang, Y. N., Zhu, C., Jiang, J., Zhang, H., Zhu, J. K., and Duan, C. G. (2020). Epigenetic regulation in plant abiotic stress responses. *J. Integr. Plant. Biol.* 62, 563–580. doi: 10.1111/jipb.12901

Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., et al. (2015). Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nat. Genet.* 47, 1149–1157. doi: 10.1038/ng.3385

Chen, D. H., Huang, Y., Jiang, C., and Si, J. P. (2018). Chromatin-based regulation of plant root development. *Front. Plant Sci.* 9:1509. doi: 10.3389/fpls.2018.01509

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chen, 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

Dahl, J. A., Jung, I., Aanes, H., Greggains, G. D., Manaf, A., Lerdrup, M., et al. (2016). Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* 537, 548–552. doi: 10.1038/nature19360

Dai, S., Zhang, Z., Chen, S., and Beachy, R. N. (2004). RF2b, a rice bZIP transcription activator, interacts with RF2a and is involved in symptom development of rice tungro disease. *Proc. Natl. Acad. Sci. U S A.* 101, 687–692. doi: 10.1073/pnas.0307687100

Deal, R. B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev. Cell* 18, 1030–1040. doi: 10.1016/j.devcel.2010.05.013

Deal, R. B., and Henikoff, S. (2011). Histone variants and modifications in plant gene regulation. *Curr. Opin. Plant Biol.* 14, 116–122. doi: 10.1016/j.pbi.2010.11.005

Deevy, O., and Bracken, A. P. (2019). PRC2 functions in development and congenital disorders. *Development* 146:dev181354. doi: 10.1242/dev.181354

Dhar, S. S., Zhao, D., Lin, T., Gu, B., Pal, K., Wu, S. J., et al. (2018). MLL4 is required to maintain broad H3K4me3 peaks and super-enhancers at tumor suppressor genes. *Mol. Cell* 70, 825–841.e6. doi: 10.1016/j.molcel.2018.04.028.

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns–from conservation to diversity. *Trends Plant Sci.* 11, 199–208. doi: 10.1016/j.tplants.2006.02.008

Gan, E. S., Xu, Y., and Ito, T. (2015). Dynamics of H3K27me3 methylation and demethylation in plant development. *Plant Signal. Behav.* 10:e1027851. doi: 10.1080/15592324.2015.1027851

Gommers, C. M., Keuskamp, D. H., Buti, S., van Veen, H., Koevoets, I. T., Reinen, E., et al. (2017). Molecular profiles of contrasting shade response strategies in wild plants: differential control of immunity and shoot elongation. *Plant Cell* 29, 331–344. doi: 10.1105/tpc.16.00790

Han, Q., Bartels, A., Cheng, X., Meyer, A., An, Y. C., Hsieh, T. F., et al. (2019). Epigenetics regulates reproductive development in plants. *Plants (Basel)* 8:564. doi: 10.3390/plants8120564

He, G., Zhu, X., Elling, A. A., Chen, L., Wang, X., Guo, L., et al. (2010). Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *Plant Cell* 22, 17–33. doi: 10.1105/tpc.109.072041

Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS One* 5:e12776. doi: 10.1371/journal.pone.0012776

Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 42, D1182–D1187. doi: 10.1093/nar/gkt1016

Kieffer, M., Master, V., Waites, R., and Davies, B. (2011). TCP14 and TCP15 affect internode length and leaf shape in *Arabidopsis*. *Plant J.* 68, 147–158. doi: 10.1111/j.1365-313X.2011.04674.x

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

Kim, J. M., Sasaki, T., Ueda, M., Sako, K., and Seki, M. (2015). Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. *Front. Plant Sci.* 6:114. doi: 10.3389/fpls.2015.00114

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868–871. doi: 10.1126/science.184.4139.868

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. doi: 10.1016/j.cell.2007.02.005

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

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, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R. J., et al. (2015). Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nat. Biotechnol.* 33, 524–530. doi: 10.1038/nbt.3208

Li, Y., Pearl, S. A., and Jackson, S. A. (2015). Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci.* 20, 664–675. doi: 10.1016/j.tplants.2015.06.013

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. *Annu. Rev. Plant Biol.* 61, 395–420. doi: 10.1146/annurev.arplant.043008.091939

Liu, H., Yu, X., Li, K., Klejnot, J., Yang, H., Lisiero, D., et al. (2008). Photoexcited CRY2 interacts with CIB1 to regulate transcription and floral initiation in *Arabidopsis*. *Science* 322, 1535–1539. doi: 10.1126/science.1163927

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., et al. (2016). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* 537, 558–562. doi: 10.1038/nature19362

Lucero, L. E., Uberti-Manassero, N. G., Arce, A. L., Colombatti, F., Alemano, S. G., and Gonzalez, D. H. (2015). TCP15 modulates cytokinin and auxin responses during gynoecium development in *Arabidopsis*. *Plant J.* 84, 267–282. doi: 10.1111/tpj.12992

Lv, J., and Chen, K. (2016). Broad H3K4me3 as a novel epigenetic signature for normal development and disease. *Genom. Proteom. Bioinf.* 14, 262–264. doi: 10.1016/j.gpb.2016.09.001

Niu, X., Luo, T., Zhao, H., Su, Y., Ji, W., and Li, H. (2020). Identification of wheat DREB genes and functional characterization of TaDREB3 in response to abiotic stresses. *Gene* 740:144514. doi: 10.1016/j.gene.2020.144514

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Probst, A. V., and Mittelsten Scheid, O. (2015). Stress-induced structural changes in plant chromatin. *Curr. Opin. Plant. Biol.* 27, 8–16. doi: 10.1016/j.pbi.2015.05.011

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Radoeva, T., Lokerse, A. S., Llavata-Peris, C. I., Wendrich, J. R., Xiang, D., Liao, C. Y., et al. (2019). A robust auxin response network controls embryo and suspensor development through a basic helix loop helix transcriptional module. *Plant Cell* 31, 52–67. doi: 10.1105/tpc.18.00518

Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., et al. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 44, W160–W165.

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

Ramirez-Prado, J. S., Piquerez, S. J. M., Bendahmane, A., Hirt, H., Raynaud, C., and Benhamed, M. (2018). Modify the histone to win the battle: chromatin dynamics in plant-pathogen interactions. *Front. Plant Sci.* 9:355. doi: 10.3389/fpls.2018.00355

Sanders, D., Qian, S., Fieweger, R., Lu, L., Dowell, J. A., Denu, J. M., et al. (2017). Histone lysine-to-methionine mutations reduce histone methylation and cause developmental pleiotropy. *Plant Physiol.* 173, 2243–2252. doi: 10.1104/pp.16.01499

Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., et al. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411. doi: 10.1038/nature01080

Sims, R. J. III, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet.* 19, 629–639. doi: 10.1016/j.tig.2003.09.007

Song, Q., Zhang, T. Z., David, M. S., 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

Stillman, B. (2018). Histone modifications: insights into their influence on gene expression. *Cell* 175, 6–9. doi: 10.1016/j.cell.2018.08.032

Tao, X., Feng, S., Zhao, T., and Guan, X. (2020). Efficient chromatin profiling of H3K4me3 modification in cotton using CUT&Tag. *Plant Methods* 16:120. doi: 10.1186/s13007-020-00664-8

Tariq, M., and Paszkowski, J. (2004). DNA and histone methylation in plants. *Trends Genet.* 20, 244–251. doi: 10.1016/j.tig.2004.04.005

Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017

Tian, L., and Chen, Z. J. (2001). Blocking histone deacetylation in *Arabidopsis* induces pleiotropic effects on plant gene regulation and development. *Proc. Natl. Acad. Sci. U S A.* 98, 200–205. doi: 10.1073/pnas.011347998

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* 45, W122–W129.

Ueda, M., and Seki, M. (2020). Histone modifications form epigenetic regulatory networks to regulate abiotic stress response. *Plant Physiol.* 182, 15–26. doi: 10.1104/pp.19.00988

Ueda, Y., Ohtsuki, N., Kadota, K., Tezuka, A., Nagano, A. J., Kadowaki, T., et al. (2020). Gene regulatory network and its constituent transcription factors that control nitrogen-deficiency responses in rice. *New Phytol.* 227, 1434–1452. doi: 10.1111/nph.16627

van Dam, S., Vosa, U., van der Graaf, A., Franke, L., and de Magalhaes, J. P. (2018). Gene co-expression analysis for functional classification and gene-disease predictions. *Brief. Bioinform.* 19, 575–592. doi: 10.1093/bib/bbw139

Van Leene, J., Blomme, J., Kulkarni, S. R., Cannoot, B., De Winne, N., Eeckhout, D., et al. (2016). Functional characterization of the *Arabidopsis* transcription factor bZIP29 reveals its role in leaf and root development. *J. Exp. Bot.* 67, 5825–5840.

Wang, K., Wang, Z., Li, F., Ye, W., Wang, J., Song, G., et al. (2012). The draft genome of a diploid cotton *Gossypium raimondii*. *Nat. Genet.* 44, 1098–1103. doi: 10.1038/ng.2371

Wang, M., Tu, L., Lin, M., Lin, Z., Wang, P., Yang, Q., et al. (2017). Asymmetric subgenome selection and cis-regulatory divergence during cotton domestication. *Nat. Genet.* 49, 579–587. doi: 10.1038/ng.3807

Wang, M., Tu, L., Yuan, D., Zhu, S., Shen, C., Li, J., et al. (2019). Reference genome sequences of two cultivated allotetraploid cottons, *Gossypium hirsutum* and *Gossypium barbadense*. *Nat. Genet.* 51, 224–229. doi: 10.1038/s41588-018-0282-x

Wang, M., Wang, P., Tu, L., Zhu, S., Zhang, L., Li, Z., et al. (2016). Multi-omics maps of cotton fibre reveal epigenetic basis for staged single-cell differentiation. *Nucleic Acids Res.* 44, 4067–4079. doi: 10.1093/nar/gkw238

Wang, X., Elling, A. A., Li, X., Li, N., Peng, Z., He, G., et al. (2009). Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. *Plant Cell* 21, 1053–1069. doi: 10.1105/tpc.109.065714

Wiles, E. T., and Selker, E. U. (2017). H3K27 methylation: a promiscuous repressive chromatin mark. *Curr. Opin. Genet. Dev.* 43, 31–37. doi: 10.1016/j.gde.2016.11.001

You, Q., Yi, X., Zhang, K., Wang, C., Ma, X., Zhang, X., et al. (2017). Genome-wide comparative analysis of H3K4me3 profiles between diploid and allotetraploid cotton to refine genome annotation. *Sci. Rep.* 7:9098. doi: 10.1038/s41598-017-09680-6

Yu, J., Jung, S., Cheng, C. H., Ficklin, S. P., Lee, T., Zheng, P., et al. (2014). CottonGen: a genomics, genetics and breeding database for cotton research. *Nucleic Acids Res.* 42, D1229–D1236. doi: 10.1093/nar/gkt1064

Zaidi, S. S., Mansoor, S., and Paterson, A. (2018). The rise of cotton genomics. *Trends Plant Sci.* 23, 953–955. doi: 10.1016/j.tplants.2018.08.009

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. *Nature* 537, 553–557. doi: 10.1038/nature19361

Zhang, T., Cooper, S., and Brockdorff, N. (2015a). The interplay of histone modifications - writers that read. *EMBO Rep.* 16, 1467–1481. doi: 10.15252/embr.201540945

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015b). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207

Zhang, W., Wu, Y., Schnable, J. C., Zeng, Z., Freeling, M., Crawford, G. E., et al. (2012). High-resolution mapping of open chromatin in the rice genome. *Genome Res.* 22, 151–162. doi: 10.1101/gr.131342.111

Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., and Jacobsen, S. E. (2009). Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* 10:R62. doi: 10.1186/gb-2009-10-6-r62

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J., et al. (2007). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol.* 5:e129. doi: 10.1371/journal.pbio.0050129

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. doi: 10.1186/gb-2008-9-9-r137

Zhang, Y., Wu, R., Qin, G., Chen, Z., Gu, H., and Qu, L. J. (2011). Over-expression of WOX1 leads to defects in meristem development and polyamine homeostasis in *Arabidopsis*. *J. Integr. Plant. Biol.* 53, 493–506. doi: 10.1111/j.1744-7909.2011.01054.x

Zheng, B., and Chen, X. (2011). Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant. Biol.* 14, 123–129. doi: 10.1016/j.pbi.2011.01.001

Zheng, D., Ye, W., Song, Q., Han, F., Zhang, T., and Chen, Z. J. (2016). Histone modifications define expression bias of homoeologous genomes in allotetraploid cotton. *Plant Physiol.* 172, 1760–1771. doi: 10.1104/pp.16.01210

Zheng, K., Wang, Y., Zhang, N., Jia, Q., Wang, X., Hou, C., et al. (2017). Involvement of PACLOBUTRAZOL RESISTANCE6/KIDARI, an atypical bHLH transcription factor, in auxin responses in *Arabidopsis*. *Front. Plant Sci.* 8:1813. doi: 10.3389/fpls.2017.01813

Zhu, L. J., Gazin, C., Lawson, N. D., Pages, H., Lin, S. M., Lapointe, D. S., et al. (2010). ChIPpeakAnno: a bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237. doi: 10.1186/1471-2105-11-237

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

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 January 2022
doi: 10.3389/fpls.2021.825810

Histone Deacetylation Controls Xylem Vessel Cell Differentiation *via* Transcriptional Regulation of a Transcription Repressor Complex OFP1/4–MYB75–KNAT7–BLH6

Risaku Hirai[1†], *Shumin Wang*[2†], *Taku Demura*[1,3,4] *and Misato Ohtani*[1,3,5*]

[1] Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan, [2] Department of Botany, University of British Columbia, Vancouver, BC, Canada, [3] RIKEN Center for Sustainable Resource Science, Yokohama, Japan, [4] Center for Digital Green-Innovation, Nara Institute of Science and Technology, Ikoma, Japan, [5] Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Japan

OPEN ACCESS

Edited by:
Nobutoshi Yamaguchi, Nara Institute of Science and Technology (NAIST), Japan

Reviewed by:
Masaru Ohme-Takagi, Saitama University, Japan
Hirokazu Tanaka, Meiji University, Japan

***Correspondence:**
Misato Ohtani
misato@edu.k.u-tokyo.ac.jp

†Present addresses:
Risaku Hirai, Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium; Center for Plant Systems Biology, Vlaams Instituut voor Biotechnologie, Ghent, Belgium
Shumin Wang, Iona Drive, Vancouver, BC, Canada

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

Received: 30 November 2021
Accepted: 28 December 2021
Published: 27 January 2022

Citation:
Hirai R, Wang S, Demura T and Ohtani M (2022) Histone Deacetylation Controls Xylem Vessel Cell Differentiation via Transcriptional Regulation of a Transcription Repressor Complex OFP1/4–MYB75–KNAT7–BLH6. Front. Plant Sci. 12:825810. doi: 10.3389/fpls.2021.825810

Xylem vessels are indispensable tissues in vascular plants that transport water and minerals. The differentiation of xylem vessel cells is characterized by secondary cell wall deposition and programmed cell death. These processes are initiated by a specific set of transcription factors, called VASCULAR-RELATED NAC-DOMAIN (VND) family proteins, through the direct and/or indirectly induction of genes required for secondary cell wall deposition and programmed cell death. In this study, we explored novel regulatory factors for xylem vessel cell differentiation in *Arabidopsis thaliana*. We tested the effects of cellular stress inducers on VND7-induced differentiation of xylem vessel cells with the *VND7–VP16–GR* system, in which VND7 activity is post-translationally induced by dexamethasone application. We established that the histone deacetylase (HDAC) inhibitors trichostatin A (TSA) and sirtinol inhibited VND7-induced xylem vessel cell differentiation. The inhibitory effects of TSA and sirtinol treatment were detected only when they were added at the same time as the dexamethasone application, suggesting that TSA and sirtinol mainly influence the early stages of xylem vessel cell differentiation. Expression analysis revealed that these HDAC inhibitors downregulated VND7-downstream genes, including both direct and indirect targets of transcriptional activation. Notably, the HDAC inhibitors upregulated the transcript levels of negative regulators of xylem vessel cells, OVATE FAMILY PROTEIN1 (OFP1), OFP4, and MYB75, which are known to form a protein complex with BEL1-LIKE HOMEODOMAIN6 (BLH6) to repress gene transcription. The KDB system, another *in vitro* induction system of ectopic xylem vessel cells, demonstrated that TSA and sirtinol also inhibited ectopic formation of xylem vessel cells, and this inhibition was partially suppressed in *knat7-1*, *bhl6-1*, *knat7-1 bhl6-1*, and quintuple *ofp1 ofp2 ofp3 ofp4 ofp5* mutants. Thus, the negative effects of HDAC inhibitors on xylem vessel cell differentiation are mediated, at least partly, by the abnormal upregulation of the transcriptional repressor complex OFP1/4–MYB75–KNAT7–BLH6. Collectively, our findings suggest that active regulation of histone deacetylation by HDACs is involved in xylem vessel cell differentiation *via* the OFP1/4–MYB75–KNAT7–BLH6 complex.

Keywords: histone deacetylation, trichostatin A, sirtinol, VND7, xylem vessel cell

INTRODUCTION

Xylem vessels are important tissues in vascular plants that transport water and minerals. During their differentiation, xylem vessel cells develop thickened secondary cell walls (SCWs) composed of cellulose, hemicellulose, and lignin, and eventually undergo programmed cell death (PCD), resulting in a hollow structure (Turner et al., 2007; Kamon and Ohtani, 2021). Advances in molecular biological research have revealed much about the molecular mechanisms of xylem vessel cell differentiation; a factor facilitating this research was the development of an artificial induction system for xylem vessel cell differentiation (Tan et al., 2019). In 2005, an artificial induction system using cultured Arabidopsis (*Arabidopsis thaliana*) cells was established, and genome-wide transcriptome data associated with xylem vessel cell differentiation was obtained with this induction system (Kubo et al., 2005). In this work, Kubo et al. (2005) identified the VASCULAR-RELATED NAC-DOMAIN (VND) family, a group of NAC-type transcription factors whose expression is upregulated in the early stages of xylem vessel cell differentiation (Kubo et al., 2005).

The Arabidopsis VND family includes VND1 to VND7, and the *VND* genes are expressed in developing xylem vessel cells (Kubo et al., 2005; Yamaguchi et al., 2008). Overexpression of *VND* genes induces ectopic SCW thickening and programmed cell death (Kubo et al., 2005; Yamaguchi et al., 2008, 2010a; Zhou et al., 2014; Endo et al., 2015), while the artificial suppression of VND function suppresses the differentiation of xylem vessel cells (Kubo et al., 2005; Yamaguchi et al., 2010b). These results suggest that the VND family functions as a master switch for xylem vessel cell differentiation. In 2010, multiple genome-wide transcriptome analyses using the *VND6* or *VND7* inducible system revealed the direct target factors of VND proteins (Ohashi-Ito et al., 2010; Zhong et al., 2010; Yamaguchi et al., 2011). The list of VND7 direct target genes included genes involved in SCW formation, such as SCW-specific cellulose synthase genes *CELLULOSE SYNTHASE A4* (*CesA4*) and *CesA7* (Brown et al., 2005); xylan synthase genes *IRREGULAR XYLEM8* (*IRX8*) and *IRX10* (Peña et al., 2007); and PCD-related protease genes *XYLEM CYSTEINE PEPTIDASE 1* (*XCP1*) (Funk et al., 2002; Avci et al., 2008) and *METACASPASE 9* (*MC9*) (Bollhöner et al., 2013). Moreover, the downstream region of VND7 contains multiple transcription factors, such as LOB DOMAIN-CONTAINING PROTEIN 30 (LBD30), LBD15, and LBD18 (Soyano et al., 2008; Ohashi-Ito et al., 2018) as well as MYB46, MYB83, and MYB63 (Ko et al., 2009, 2012; Zhou et al., 2009; Zhong and Ye, 2012). The LBD proteins positively regulate *VND7* expression, suggesting the existence of positive feedback regulation between VND7 and LBD (Soyano et al., 2008; Ohashi-Ito et al., 2018). MYB46 and MYB83 induce the expression of SCW-related genes such as *CesA*, which is also targeted by VND7 (Ko et al., 2012; Zhong and Ye, 2012), and form a so-called feed-forward loop with VND7 (Taylor-Teeples et al., 2015; Turco et al., 2019).

While the VND family genes positively regulate xylem vessel cell differentiation, other transcription factors inhibit xylem vessel cell differentiation, including VND-INTERACTING2 (VNI2) (Yamaguchi et al., 2010b); XYLEM NAC DOMAIN1 (XND1) (Zhao et al., 2007; Zhang et al., 2020); the homeobox transcription factors BEL1-LIKE HOMEODOMAIN6 (BLH6; Liu et al., 2014; Liu and Douglas, 2015), KNOTTED ARABIDOPSIS THALIANA3 (KNAT3), and KNAT7 (Bhargava et al., 2010; Li et al., 2011, 2012; Liu and Douglas, 2015; Wang et al., 2020); OVATE FAMILY PROTEIN1 (OFP1) and OFP4 (Li et al., 2011; Liu and Douglas, 2015); and the MYB transcription factors MYB4, MYB5, MYB7, MYB32, and MYB75 (Preston et al., 2004; Ko et al., 2009; Bhargava et al., 2010, 2013). Among these, VNI2 and XND1, which are NAC-type transcription factors, have been suggested to physically interact with VND7 to form a heterodimer that represses VND7 function (Yamaguchi et al., 2010b; Zhang et al., 2020).

The above findings suggest the existence of a complex transcriptional regulatory network consisting of VND7-based positive regulation and negative regulation by multiple classes of transcription factors (Ohashi-Ito and Fukuda, 2010; Hussey et al., 2013; Nakano et al., 2015; McCahill and Hazen, 2019; Ohtani and Demura, 2019). A gene co-expression analysis suggested that the expression and interaction patterns of these transcription factor genes, including the VND family members, could be different in the presence and absence of stresses such as salt and drought (Taylor-Teeples et al., 2015; Ohtani and Demura, 2019). Molecular genetics research also revealed that the activity of VND family proteins can be actively regulated in response to light (Tan et al., 2018) and cellular thiol conditions (Kawabe et al., 2018; Ohtani et al., 2018). Thus, it is highly possible that stress can modify the transcriptional regulatory network, especially the balance between transcriptionally positive and negative regulation modules, for proper xylem vessel cell differentiation.

To obtain clues into novel regulatory factors involved in such modification, we tested the effects of cellular stresses on xylem vessel cell differentiation initiated by VND7. We used Arabidopsis *VND7–VP16–GR* plants overexpressing a chimeric protein of VND7, a transcriptional activation domain VP16, and the glucocorticoid receptor (GR) (Yamaguchi et al., 2010a). In this system, VND7 activity is post-translationally activated by treatment with glucocorticoids, such as dexamethasone (DEX), leading to ectopic induction of xylem vessel cell differentiation in a DEX concentration-dependent manner (Yamaguchi et al., 2010a; Hirai et al., 2019; **Supplementary Figure 1A**). We treated the Arabidopsis *VND7–VP16–GR* seedlings with known cellular stress inducers and found that the histone deacetylase (HDAC) inhibitors trichostatin A (TSA) and sirtinol significantly inhibited ectopic xylem vessel cell differentiation. Further expression analysis demonstrated that the HDAC inhibitor treatment upregulated negative regulators of xylem vessel cell differentiation, such as OVATE FAMILY PROTEIN1 (OFP1), OFP4, and MYB75, which form a protein complex with BEL1-LIKE HOMEODOMAIN6 (BLH6) to repress gene transcription. Indeed, the *knat7* and *ofp* mutations suppressed the inhibitory effects of the HDAC inhibitor on xylem vessel cell differentiation. These results suggest that HDACs play important roles in xylem vessel cell differentiation through the regulation of a transcriptional repression complex.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

The Arabidopsis (*A. thaliana*) *VND7–VP16–GR* line (Col-0) was described in Yamaguchi et al. (2010a), and *ofp1*, *ofp4*, *knat7-1*, *blh6-1*, and *knat7-1 blh6-1* were reported in Wang et al. (2007), Li et al. (2011), and Liu et al. (2014). T-DNA insertion lines for *ofp2* (SALK_122550, with the insertion at the 3′UTR; **Supplementary Figure 2**), *ofp3* (GABI_167F01, with the insertion at the exon), and *ofp5* (SALK_203823, with the insertion at the exon) were obtained from the Arabidopsis Biological Resource Center (ABRC). Quintuple *ofp1 ofp2 ofp3 ofp4 ofp5* mutants were generated by genetic crossing, and the genotypes were confirmed by PCR. The seedlings were grown on Murashige and Skoog (MS) medium (Wako, Japan) containing 1% (w/v) sucrose (Nacalai tesque, Japan), 0.05% (w/v) MES (nacalai tesque), and 0.6% (w/v) Gellan gum (Wako), adjusted to pH 5.7 under continuous light at 22°C.

DEX Treatment and Cellular Stress Inducer Treatment

Seven-day-old seedlings of the *VND7–VP16–GR* line were treated with 10 nM DEX (Sigma-Aldrich, St. Louis, MO, USA), as previously described in Hirai et al. (2019). Briefly, the seedlings were transferred into a 12-well plate (Corning) containing the 10 nM DEX solution and incubated for 3 d. For the treatment with cellular stress inducers (**Supplementary Table 1**), the inducers were added to the 10 nM DEX solution with the final concentration described in **Supplementary Table 1**; in addition to the reported concentration that can affect plant cell activity (**Supplementary Table 1**), one-tenth and 10-fold concentrations were tested for each chemical. To check the involvement of protein *S*-nitrosylation, or auxin response, in the downstream of HDAC inhibitors, the *VND7–VP16–GR* seedlings were treated by 10 nM DEX and 5 μM TSA, in the presence or absence of 500 μM 2-(4-carboxyphenyl)-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide (cPTIO), a nitric oxide scavenger, or by 10 nM DEX, in the presence or absence of 10 μM 1-naphthaleneacetic acid (NAA), for 3 d, respectively. To examine the time window where HDAC inhibitors can influence xylem vessel cell differentiation, TSA and sirtinol were added at 0, 6, and 12 h of DEX treatment.

KDB Treatment

KBD treatment was carried out as described previously (Tan et al., 2018, 2019). The cotyledons of 7-day-old seedlings were excised and incubated in half-strength MS liquid medium supplemented with KDB hormone mixture (50 ng/ml kinetin, 500 ng/ml 2,4-dichlorophenoxyacetic acid, and 1 μm brassinolide) at 22°C under continuous light for 4 d.

Microscopy Observation

The SCW deposition ratio in the *VND7–VP16–GR* leaves was determined according to the method in Hirai et al. (2019). Cotyledon samples were fixed, stained with propidium iodide (PI), and mounted using TOMEI-I (Hasegawa et al., 2016; Hirai et al., 2019). These samples were observed with a FV-10i confocal microscope (Olympus). The confocal images were processed using ImageJ (https://imagej.nih.gov/ij/index.html) with the plugin MosaicJ (Thévenaz and Unser, 2007) to obtain maximum intensity projection images of whole cotyledons. We measured the area with SCW-signals manually and calculated the ratio of SCW-positive cell regions.

To observe KDB-treated cotyledons, the cotyledon samples were fixed with a mixture of 10% (v/v) acetic acid and 90% (v/v) ethanol. The samples were then hydrated with a 90, 70, 50, and 30% (v/v) ethanol series for 20 min each and then transferred into distilled water. Finally, the samples were mounted in clearing solution (chloral hydrate:water:glycerol 8:1:2 [w/v/v]). Images were taken using a light microscope (BX53; Olympus) equipped with differential interference contrast (DIC) optics and a digital camera (DP72; Olympus).

Reverse Transcription Quantitative PCR Analysis

Reverse transcription quantitative Quantitative RT-PCR analysis was performed as described in Yamaguchi et al. (2011) and Hirai et al. (2019). The seedlings were collected at 0, 6, 12, 18, 24, 48, and 72 h after DEX treatment and then ground in liquid nitrogen with a TissueLyser II (Qiagen). Total RNA was isolated using an RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. cDNA synthesis was performed with 1 μg of total RNA using a Transcriptor First-Strand cDNA Synthesis Kit (Roche). Quantitative PCR was carried out with LightCycler 480 DNA SYBR Green (Roche). The expression levels of tested genes were normalized with that of *UBQ10*. The primers used in this study are shown in **Supplementary Table 2**.

Statistical Analysis

For the statistical analysis of ratio of SCW positive regions and gene expression levels, we performed Student's *t*-test between the mock-treated control and the samples treated with TSA or sirtinol. In the case of multiple comparison test, we performed Tukey's test (Tukey HSD, honestly significant difference). All tests were performed with Microsoft Excel software (ver 16.55, Microsoft) or in R (https://www.R-project.org).

RESULTS AND DISCUSSION

HDAC Inhibitors Inhibit VND7-Induced Xylem Vessel Cell Differentiation

To identify novel factors affecting VND7-induced xylem vessel cell differentiation, we examined the effects of known cellular stress inducers (50 chemicals; **Supplementary Table 1**) on ectopic differentiation of xylem vessel cells in *VND7–VP16–GR* (**Supplementary Figure 1A**). First, we screened the cellular stress inducers by determining if they inhibited bleaching of cotyledons, which is considered to reflect PCD progression (Yamaguchi et al., 2010a; Hirai et al., 2019; **Supplementary Figure 1B**). After the first screening, 12 chemicals were found to inhibit the bleaching of cotyledons (**Supplementary Figure 1B**). As a second screening, the treated cotyledons were examined for SCW deposition (**Supplementary Figure 1B**). Of the 12 chemicals, five (i.e.,

citrate acid, oxidized glutathione (GSSG, glutathione-S-S-glutathione), reduced glutathione (GSH, glutathione-SH), trichostatin A (TSA), and sirtinol), reduced the ratio of SCW deposition significantly (**Supplementary Figure 3**).

The *S*-nitrosylation of VND7 is important for the regulation of VND7 transcriptional activity (Kawabe et al., 2018; Ohtani et al., 2018). GSSG and GSH can affect cellular thiol environments, possibly leading to the disturbance of protein *S*-nitrosylation regulation (Meyer and Hell, 2005). Thus, the inhibitory effects of GSSG and GSH on xylem vessel cell differentiation in *VND7–VP16–GR* may reflect such functional disruption of VND7 by the disturbance of cellular thiol homeostasis. Interestingly, TSA and sirtinol, which are histone deacetylase (HDAC) inhibitors (Grozinger et al., 2001; Chang and Pikaard, 2005; Bourque et al., 2011; Liu et al., 2017; Mengel et al., 2017; Ueda et al., 2017), were among the five chemicals that substantially reduced the ratio of SCW deposition (**Supplementary Figure 3**). TSA and sirtinol are known to inhibit class I/II and class III HDACs, respectively (Yoshida and Horinouchi, 1999; Grozinger et al., 2001). Our detailed analysis showed that more than 0.5 μM TSA and 10 μM sirtinol significantly inhibited the bleaching of seedlings (**Figures 1A–G**) as well as SCW deposition in cotyledons (**Figures 1E,H**). These observations indicate that HDAC activity is important for VND7-induced xylem vessel cell differentiation.

TSA can also promote protein *S*-nitrosylation (Mengel et al., 2017), and sirtinol can induce auxin response (Zhao et al., 2003). Therefore, we tested whether the negative effect of TSA and sirtinol on xylem vessel cell differentiation depends on protein *S*-nitrosylation or auxin response. For this, we co-treated *VND7–VP16–GR* seedlings with the nitric oxide scavenger 2-(4-carboxyphenyl)-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide (cPTIO), which inhibits protein *S*-nitrosylation (Mengel et al., 2017; Ageeva-Kieferle et al., 2021), or with synthetic auxin 1-naphthaleneacetic acid (NAA). The inhibition of xylem vessel cell differentiation by TSA was not affected by the additional application of cPTIO (**Supplementary Figure 4**), and NAA did not inhibit the differentiation (**Supplementary Figure 5**), suggesting that the inhibitory effect on VND7-induced xylem vessel cell differentiation was not due to protein *S*-nitrosylation nor auxin response.

HDAC Inhibitors Disturb the Early Stages of Xylem Vessel Cell Differentiation

Next, to examine the effect of application timing of HDAC inhibitors on xylem vessel cell differentiation, TSA and sirtinol were added to the *VND7–VP16–GR* seedlings at 0, 6, and 12 h of DEX treatment. We found a significant reduction in the degree of bleaching and SCW deposition when TSA or sirtinol were added at 0 h of DEX treatment (**Figures 2A–C,J**). By contrast, when we added TSA or sirtinol at 6 or 12 h of DEX treatment, no significant difference was detected in the degree of bleaching and SCW deposition, although the addition of HDAC inhibitors at 6 h of DEX treatment slightly decreased the degree of SCW deposition (**Figures 2D–J**). This clearly demonstrated that HDAC inhibition affected the molecular processes at the early stages of xylem vessel cell differentiation, which occur up to 6 h after DEX treatment.

HDAC Inhibitors Suppress the Upregulation of VND7-Downstream Genes, Especially at Early Stages of Xylem Vessel Cell Differentiation

Our observations indicated that the HDAC inhibitors inhibited both SCW deposition and PCD progression (**Figures 1**, **2**). Therefore, we further determined the effects of TSA and sirtinol on the expression of VND7 downstream genes (**Figure 3** and **Supplementary Figure 6**). Seven-day-old *VND7–VP16–GR* seedlings were treated with the 10 nM DEX solution with or without TSA or sirtinol. Quantitative RT-PCR analysis was performed for VND7 downstream genes, such as the PCD-related genes *XCP1* and *MC9*; transcription factor genes *LBD30*, *MYB46*, and *MYB63*; cellulose and xylan biosynthesis genes *CESA7/IRX3* and *IRX8*, respectively; and a lignin biosynthesis gene *CAFFEOYL COENZYME A ESTER O-METHYLTRANSFERASE7* (*CCoAOMT7*). In addition, we examined the expression of endogenous *VND7*. The upregulation of these VND7 downstream genes was basically repressed by treatment with TSA or sirtinol (**Figure 3** and **Supplementary Figure 6**). Gene upregulation that occurred within 12 h of DEX treatment was strongly inhibited (**Figure 3** and **Supplementary Figure 6**), in accordance with the results shown in **Figure 2**. Collectively, these results indicate that the HDAC inhibitors affected the molecular events at the very early stages of xylem vessel cell differentiation, possibly inhibiting VND7 activity itself or processes close to the transactivation of gene expression by VND7.

Previously, a DEX concentration-dependent reduction in xylem vessel cell differentiation was observed in *VND7–VP16–GR* seedlings (Hirai et al., 2019). This indicates that the degree of VND7 activity, that is, the transactivation activity of downstream genes, is a crucial factor determining the progression of xylem vessel cell differentiation (Hirai et al., 2019). To clarify how TSA and sirtinol affect VND7-based transcriptional regulation for xylem vessel cell differentiation, we performed a hierarchical clustering analysis of VND7 downstream genes based on their expression patterns (shown in **Figure 3** and **Supplementary Figure 6**). The published data showed that *LBD30*, *MYB46*, *XCP1*, and *MC9* are direct targets of VND7 (Ohashi-Ito et al., 2010; Zhong et al., 2010; Yamaguchi et al., 2011), whereas *MYB63*, *CESA7*, *IRX8*, and *CCoAOMT7* are direct targets of MYB46 (Ko et al., 2009; Kim et al., 2012; Zhong and Ye, 2012). The genes could be separated into three groups based on their DEX concentration-dependent expression patterns. First group contained *MYB46* and its direct targets (*CESA7*, *IRX8*, and *CCoAOMT7*), second group contained *VND7* and two of its direct targets (*LBD30* and *XCP1*), and third one was *MYB63* (**Figure 4A**), reflecting the transcriptional hierarchy of VND7 and MYB46 (**Figure 4A**; Hirai et al., 2019). However, the expression patterns in TSA- and sirtinol-treated samples differently grouped these genes (**Figures 4B,C**). This suggests that the HDAC inhibitors do not simply repress VND7

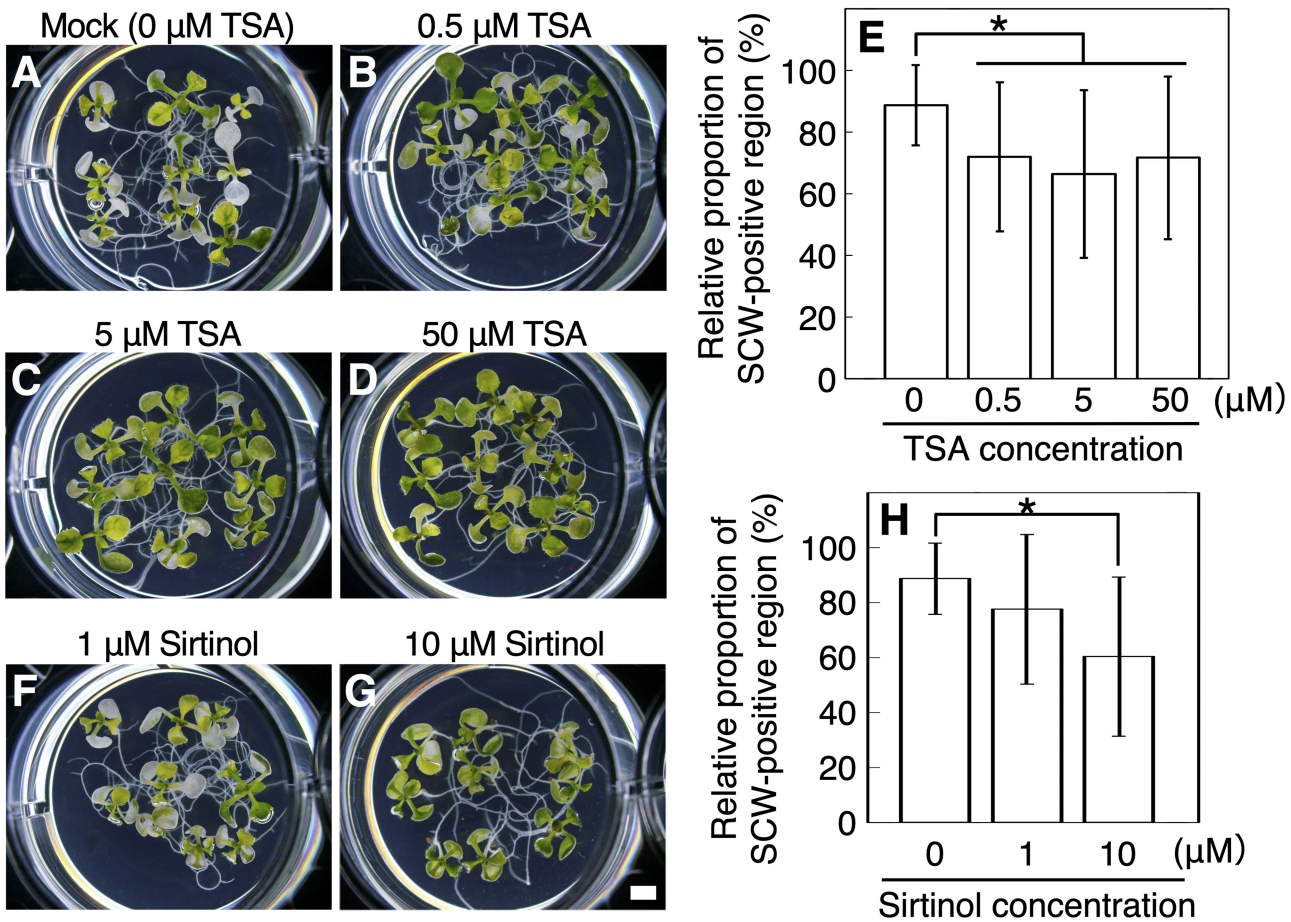

FIGURE 1 | Histone deacetylase inhibitors inhibit VND7-based xylem vessel cell differentiation. **(A–E)** Seven-day-old *VND7–VP16–GR* seedlings were soaked in the distilled water with 10 nM dexamethasone (DEX) and 0 μM **(A)**, 0.5 μM **(B)**, 5 μM **(C)**, or 50 μM **(D)** trichostatin A (TSA) for 3 d. **(F–H)** The *VND7–VP16–GR* seedlings were treated with 10 nM DEX and 1 μM **(F)** or 10 μM **(G)** sirtinol, and incubated for 3 d. **(E,H)** Ectopic deposition of secondary cell walls (SCWs) in cotyledons was observed with a confocal microscope after propidium iodide staining, and the SCW-positive regions were measured in each cotyledon. The relative proportions of SCW-positive cell regions were calculated manually using an image analysis performed in ImageJ. Results are shown as means ± SD ($n > 10$). Asterisks indicate statistically significant differences compared with the mock control (Student's *t*-test, $p < 0.05$). Bar, 1 mm.

activity as the decrease in DEX concentration does, but possibly change the regulatory relationship between the transcription factor and targets. Notably, the cluster structures were similar between the TSA- and sirtinol-treated samples (**Figures 4B,C**). Therefore, both class I/II and class III HDAC activities might be important for maintaining the transcriptional hierarchy of VND7 and MYB46 for the proper progression of xylem vessel cell differentiation.

HDAC Inhibitors Inhibit Xylem Vessel Cell Differentiation Through the Upregulation of the OFP1/4–MYB75–KNAT7–BLH6 Transcriptional Repression Complex

The inhibition of HDACs should lead to histone hyperacetylation, resulting in an increase in gene expression. Therefore, we hypothesized that the HDAC inhibitors would upregulate the expression of negative regulators of xylem vessel cell differentiation. The expression patterns of genes encoding 11 well-known transcription factors functioning as negative regulators of xylem vessel cell differentiation, namely *VNI2*, *KNAT7*, *XND1*, *BLH6*, *OFP1*, *OFP4*, *MYB4*, *MYB5*, *MYB7*, *MYB32*, and *MYB75*, were assessed (**Figure 5** and **Supplementary Figures 7, 8**). Among these negative regulators, *KNAT7*, *OFP1*, *OFP4*, and *MYB75* were differentially expressed under the HDAC inhibitor treatment; the TSA treatment upregulated *OFP1* and *MYB75*, whereas sirtinol increased *OFP4* expression, within 12 h of DEX treatment (**Figure 5**). Both inhibitors also decreased the expression levels of *KNAT7* after 24 h of DEX treatment. Since the HDAC inhibitors would disturb early stages of xylem vessel cell differentiation, i.e., within 6 h after the DEX treatment (**Figure 2**), the inhibitory effects of HDAC inhibitors on xylem vessel cell differentiation can be attributed to the upregulation of *OPF1*, *OFP4*, and *MYB75* at the early stages of xylem vessel cell differentiation. OPF1, OFP4, and MYB75 form a transcriptional repression complex with

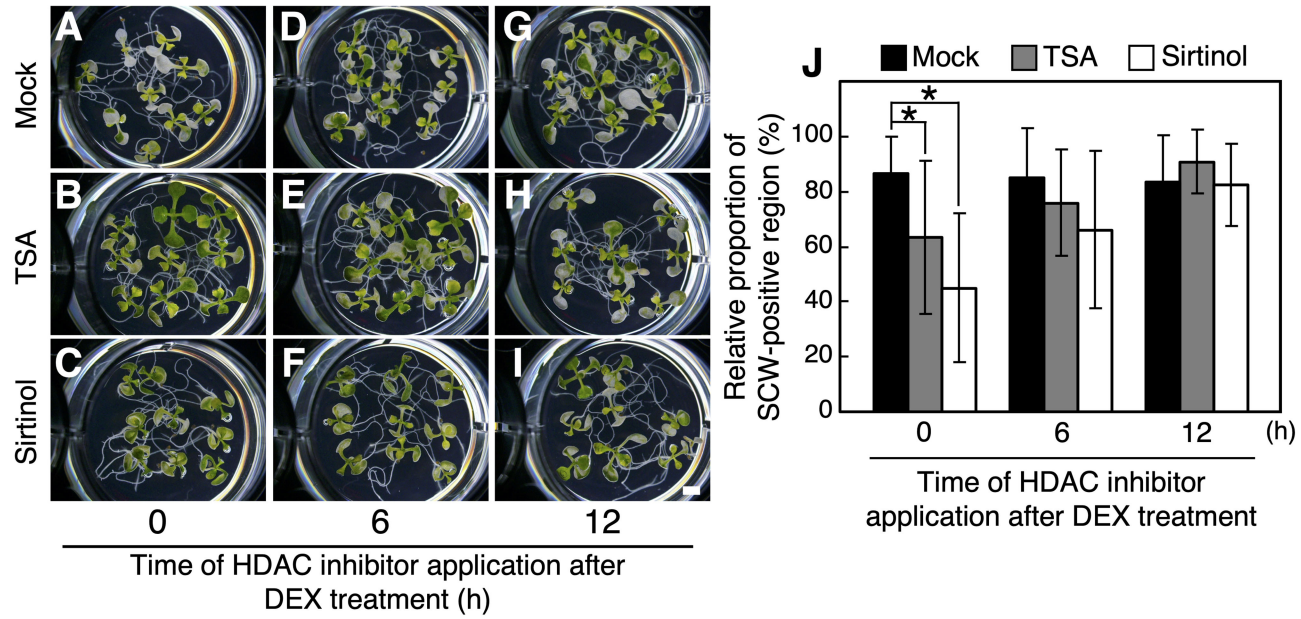

FIGURE 2 | Effects of application time of histone deacetylase inhibitors on VND7-based xylem vessel cell differentiation. Seven-day-old *VND7–VP16–GR* seedlings were treated with 10 nM dexamethasone (DEX) without any inhibitors **(A–C)**, with 5 μM trichostatin A (TSA) **(D–F)**, or with 10 μM sirtinol **(G–I)** for 3 d. TSA or sirtinol was added at 0 h **(A,D,G)**, 6 h **(B,E,H)**, and 12 h of DEX treatment. **(J)** Ectopic deposition of secondary cell walls (SCWs) in cotyledons was observed with a confocal microscope after propidium iodide staining, and the SCW-positive regions were measured in each cotyledon. The relative proportions of SCW-positive cell regions were calculated manually using an image analysis performed in ImageJ. Results are shown as means ± SD ($n > 10$). Asterisks indicate statistically significant differences compare with the mock control (Student's *t*-test, $p < 0.05$). Bar, 1 mm.

KNAT7 and BLH6 to repress the expression of genes required for xylem cell differentiation (Bhargava et al., 2010; Li et al., 2011, 2012; Liu et al., 2014; Liu and Douglas, 2015; Wang et al., 2020). Thus, these findings strongly suggest that the HDAC inhibitors inhibit xylem cell differentiation through an increase in the activity of the OFP1/4–MYB75–KNAT7–BLH6 transcriptional repression complex.

To test this possibility, we performed a mutant analysis with *knat7-1*, *blh6-1*, *knat7-1 blh6-1*, and quintuple *ofp1 ofp2 ofp3 ofp4 ofp5* mutants with a different induction system for xylem cell differentiation, the KDB system (Tan et al., 2018; **Figure 6**). *OFP1*, *OFP2*, *OFP3*, *OFP4*, and *OFP5* are close homologs in Arabidopsis. Therefore, as expected, single *ofp* mutants did not show any obvious phenotypic differences in vascular bundles compared with the wild type (**Supplementary Figure 9**). Thus, we newly established the quintuple *ofp1 ofp2 ofp3 ofp4 ofp5* mutants. In the KDB system, phytohormone treatment can induce ectopic xylem vessel cells (Tan et al., 2018, 2019). In the wild type, we recognized two types of ectopic xylem vessel cells: ectopic xylem vessel cells transdifferentiated from mesophyll cells (indicated by white arrows in **Figure 6A**) and ectopic xylem vessel cells around endogenous xylem vessels, probably originating from vascular cells (indicated by yellow triangles in **Figure 6A**). Interestingly, TSA and sirtinol strongly inhibited the transdifferentiation of ectopic xylem vessel cells from mesophyll cells (**Figure 6A**). HDAC inhibitor treatment significantly reduced the total number of ectopic xylem vessel cells (**Figure 6B**). We then checked the *knat7-1*, *blh6-1*, *knat7-1 blh6-1*, and *ofp1 ofp2 ofp3 ofp4 ofp5* mutants for ectopic xylem vessel cell differentiation in the presence or absence of TSA treatment (**Figures 6C–E**). Under the mock-treated condition, none of the mutants differed significantly from the wild type with respect to their efficiency of ectopic xylem vessel cell differentiation (**Figure 6C** and **Supplementary Figure 10**). However, TSA treatment significantly increased the number of ectopic xylem vessel cells in the mutants compared with the wild type (**Figure 6E**), clearly indicating the involvement of KNAT7, BLH6, and OFPs in the negative effects of TSA on ectopic xylem vessel cell differentiation, as expected. However, any tested mutations could not recover the transdifferentiation of mesophyll cells into xylem vessel cells after the TSA treatment (**Figure 6C**). Instead, we recognized the mutant-specific types of ectopic xylem vessel cells, which were located near endogenous xylem vessels, but their origin did not appear to be vascular cells based on their cell shapes (indicated by red triangles in **Figures 6C,D**). These observations collectively suggest that the OFP1/4–MYB75–KNAT7–BLH6 transcriptional repression complex is involved in the inhibition of xylem vessel cell differentiation in the cells near vascular tissues, but not in mesophyll cells, under treatment with HDAC inhibitors.

It has been shown that the transdifferentiation of mesophyll cells into xylem vessel cells is required cell dedifferentiation and vascular stem cell formation (Kondo et al., 2016; Tan et al., 2018; Furuya et al., 2021), which should contain multiple

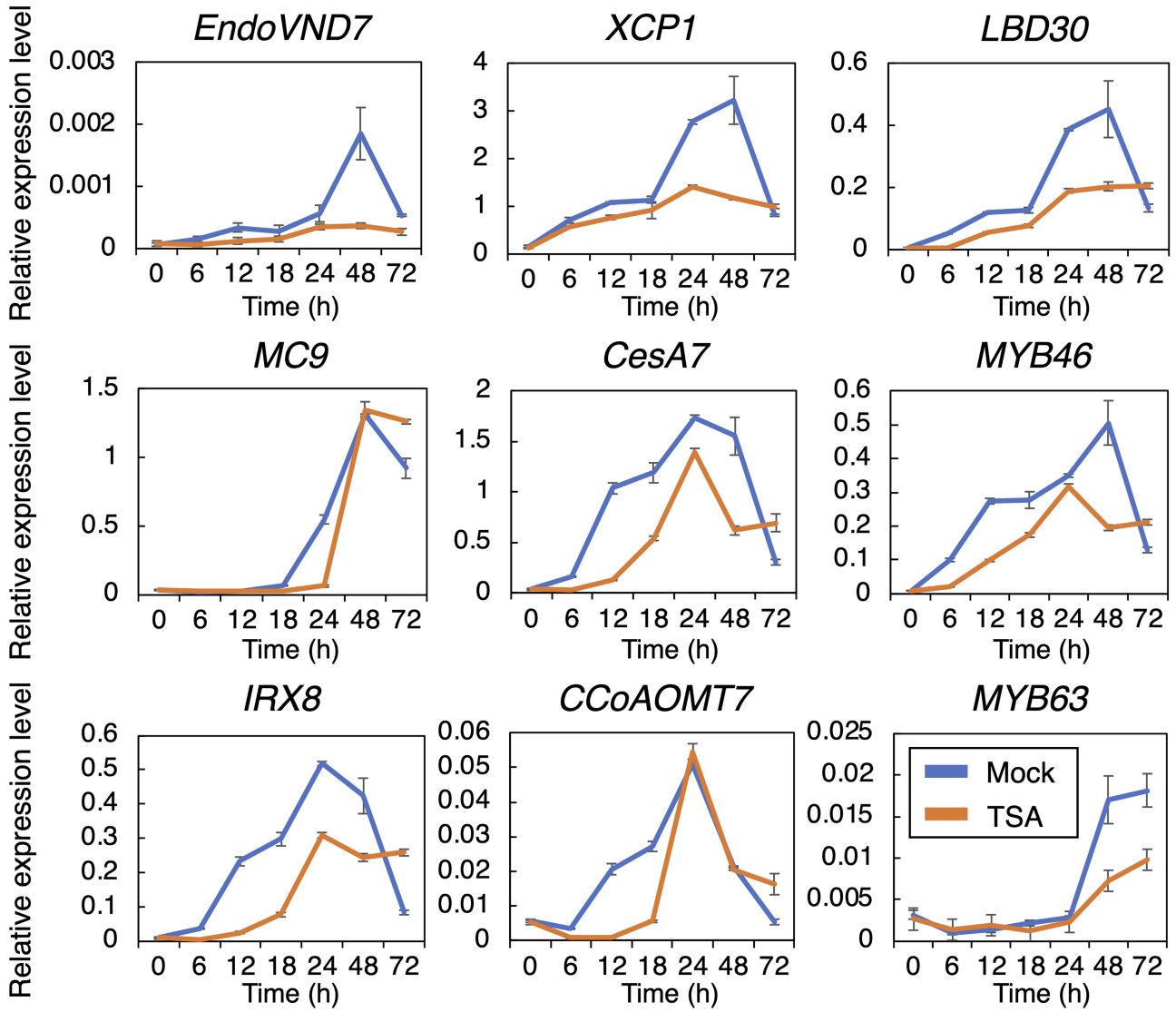

FIGURE 3 | Reverse transcription quantitative PCR analysis of the genes downstream of VND7 in *VND7–VP16–GR* seedlings treated with dexamethasone (DEX) and trichostatin A (TSA). Seven-day-old *VND7–VP16–GR* seedlings were treated with 10 nM DEX and 5 μM TSA and sampled after 0, 6, 12, 18, 24, 48, and 72 h. The expression levels of the genes downstream of VND7 were normalized to the expression level of the internal control *UBIQUITIN10*. Results are shown as means ± SD ($n = 3$).

factors regulated by HDACs. In contrast, the cells near vascular tissues, such as vascular bundle sheath cells, are known to be differentiated into xylem vessels in response to stresses (Reusche et al., 2012), suggesting that they should keep certain levels of competency for xylem vessel cell differentiation originally. Therefore, the HDAC regulatory targets would be limited to the OFP1/4–MYB75–KNAT7–BLH6 complex in the cells near vascular tissues. Moreover, considering the fact that the OFP1/4–MYB75–KNAT7–BLH6 complex can repress the gene transcription by transactivation domain such as VP16 in the transient expression assay (Li et al., 2011, 2012; Liu et al., 2014; Liu and Douglas, 2015; Wang et al., 2020), it is highly possible that this complex would directly repress the transactivation of genes for xylem vessel cell differentiation by VND7, MYB46 and/or MYB83, which are transcriptional activators. Further analysis will reveal the details of molecular mechanisms for the OFP1/4–MYB75–KNAT7–BLH6 complex-based inhibition of xylem vessel cell differentiation.

CONCLUSION AND PERSPECTIVES

In the current work, we demonstrated the roles of histone deacetylation in regulating the OFP1/4–MYB75–KNAT7–BLH6 transcriptional repression complex (Li et al., 2011, 2012; Liu et al., 2014; Liu and Douglas, 2015; Wang et al., 2020) during xylem vessel cell differentiation (**Supplementary Figure 11**). The

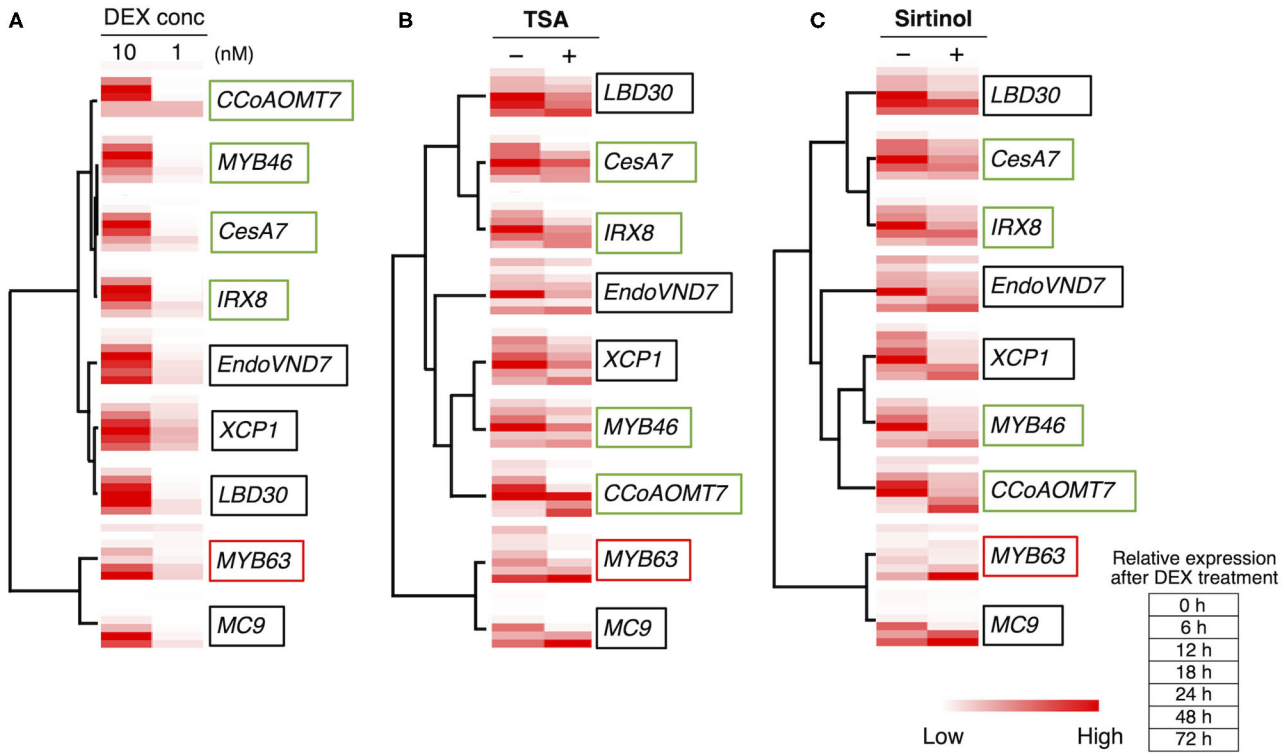

FIGURE 4 | Hierarchical clustering analysis of the expression patterns of the genes downstream of VND7. The hierarchical clustering analysis was performed using the normalized quantitative RT-PCR data presented in Hirai et al. (2019) **(A)**, in **Figure 3 (B)**, and in **Supplementary Figure 3 (C)**. Each column shows the relative expression level of each gene at each time point of DEX treatment by color shading.

Arabidopsis genome harbors 22 genes that encode HDAC proteins, which are classified into three groups, REDUCED POTASSIUM DEFICIENCY 3 (RPD3)-like HDACs (class I; 16 genes), HD-tuins (class II; 4 genes), and sirtuins (class III; 2 genes) (Hollender and Liu, 2008). The knockout mutant of *HDT1*, one of HD-tuins type HDAC genes, produced the decreased size of xylem vessels with enhanced SCW thickness (Zhang et al., 2019), suggesting a relationship between xylem vessel cell differentiation and HD-tuins type HDACs. In addition, *OFP1* and *MYB75* expression has been reported to be increased in HDAC mutants; *OFP1* and *MYB75* are upregulated in *srt1 srt2* and in *hda19*, respectively (Zhang et al., 2018; Ning et al., 2019). Moreover, the histone acetylation level was increased at the *OFP1* gene locus in *hda6* and in the *MYB75* gene locus in *hda19* (Ning et al., 2019; Ageeva-Kieferle et al., 2021), suggesting that *OFP1* and *MYB75* are targets of histone acetylation-based active regulation of gene expression. Further analysis of the contribution of HDA proteins to the regulation of *OFP1* and *MYB75* expression will provide insight into how the OFP1/4–MYB75–KNAT7–BLH6 transcriptional repression complex affects xylem vessel cell differentiation.

Previous studies have shown that xylem vessel cell differentiation is affected by a variety of environmental stresses, such as wounding (Jacobs, 1952; Comer, 1978), salt stress (Hilal et al., 1998; Taylor-Teeples et al., 2015), bacterial infection (Reusche et al., 2012), and the light environment (Tan et al., 2018). Moreover, glutathione (Henmi et al., 2001, 2005) and nitric oxide (NO) (Kawabe et al., 2018; Ohtani et al., 2018) have been reported to affect xylem vessel cell differentiation. In particular, VND7 activity is affected by *S*-nitrosoglutathione (GSNO) application and VND7 can be *S*-nitrosylated to regulate the transcriptional activity of VND7 (Kawabe et al., 2018; Ohtani et al., 2018). Accordingly, we identified GSSG and GSH, important regulators of cellular thiol homeostasis, as inhibitors of VND7-based xylem vessel cell differentiation (**Supplementary Figure 3**). Recently, it has been reported that HDA6 is *S*-nitrosylated in response to NO *via* GSNO metabolism (Ageeva-Kieferle et al., 2021). This observation strongly suggests that HDAC, a crucial epigenomic regulator linking stress responses and gene expression (Mengel et al., 2017; Song et al., 2017; Ueda et al., 2017), is a key modulator of VND7-based transcriptional switching for xylem vessel cell differentiation (**Supplementary Figure 11**).

In summary, we suggest a novel environmental response strategy in plants, in which xylem vessel cell differentiation is regulated to match xylem vessel activities with environmental conditions. HDACs are a key part of this process, regulating histone deacetylation at the *OFP1/4* and *MYB75* gene loci (**Supplementary Figure 9**). The involvement of histone

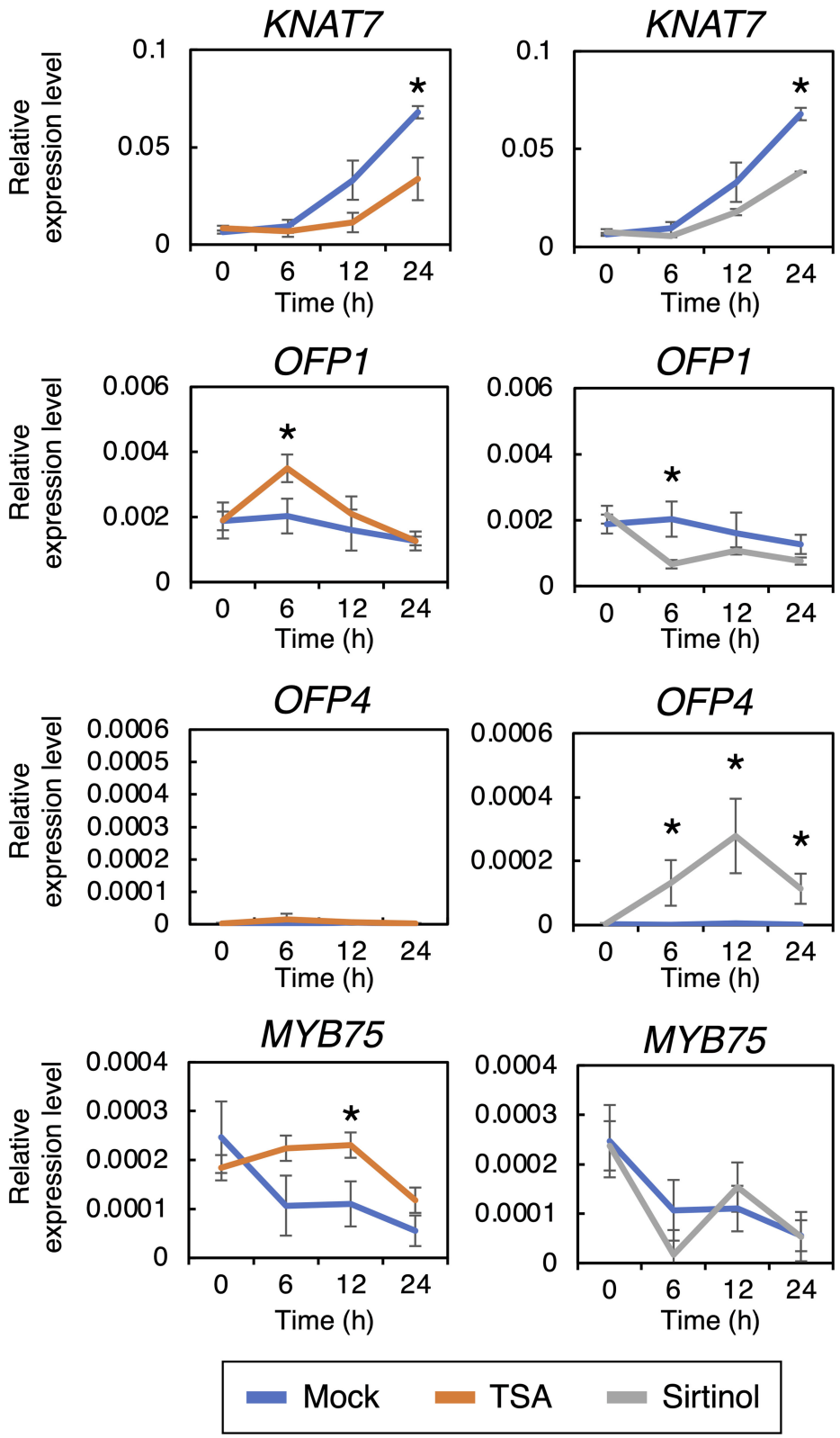

FIGURE 5 | Reverse transcription quantitative PCR analysis of genes encoding transcription factors known to negatively regulate xylem vessel cell differentiation. Seven-day-old *VND7–VP16–GR* seedlings were treated with 10 nM dexamethasone (DEX) and 5 μM trichostatin A TSA or 10 μM sirtinol and sampled after 0, 6, 12, 18, and 24 h. The transcript levels of the genes downstream of VND7 were quantified by reverse transcription quantitative PCR and normalized to the expression level of the internal control *UBIQUITIN10*. Results are shown as means ± SD ($n = 3$). Asterisks indicate statistically significant differences compared with the mock control (Student's *t*-test, $p < 0.05$).

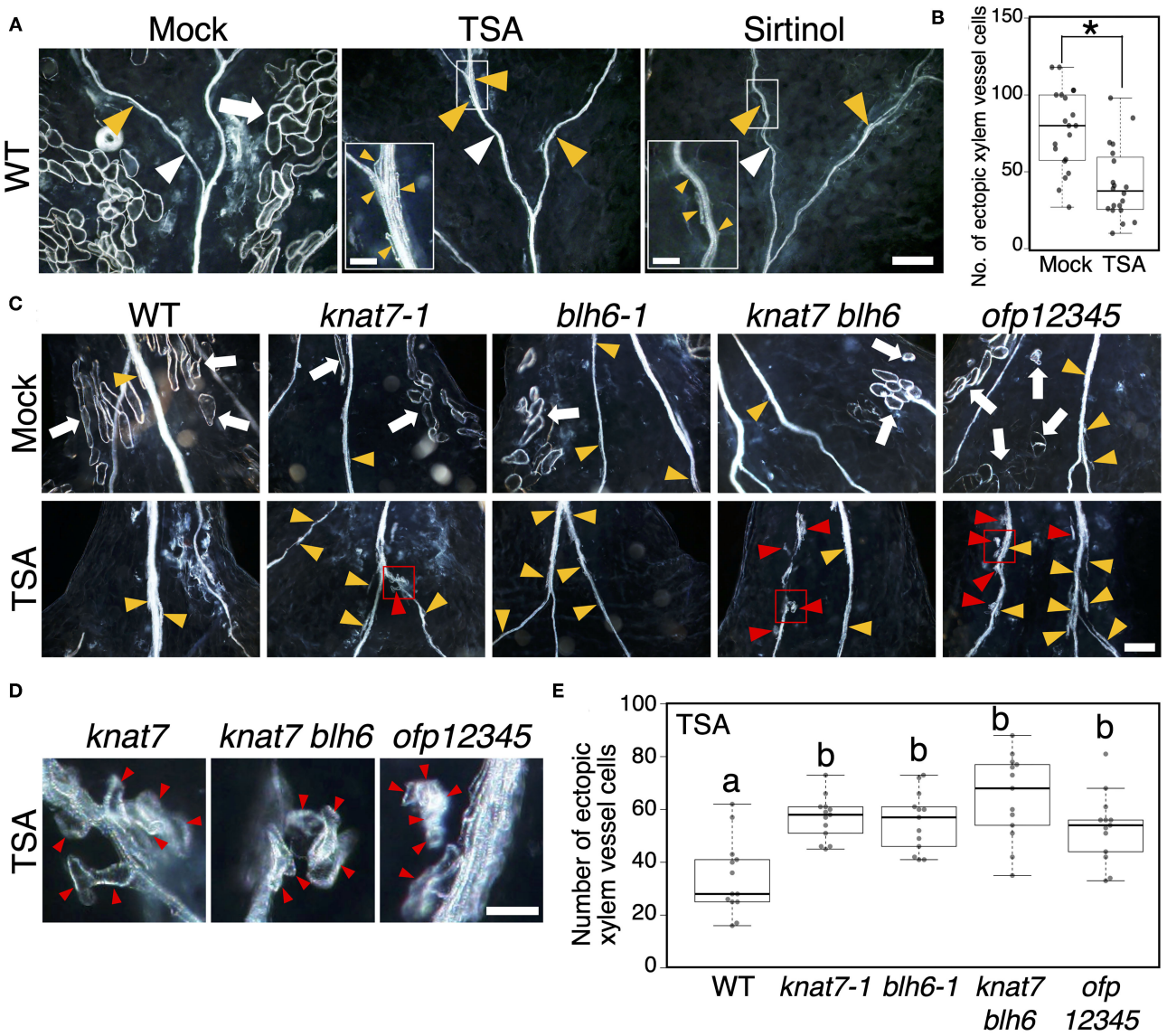

FIGURE 6 | Histone deacetylase inhibitors disturb ectopic xylem vessel cell differentiation in the KDB system. **(A)** Cotyledons were cut from seven-day-old wild-type seedlings, and ectopic xylem vessel cell differentiation was induced by phytohormonal treatment (the KDB system: Tan et al., 2018), with dimethyl sulfoxide(DMSO) (mock control), 5 μM trichostatin A (TSA), or 10 μM sirtinol, for 5 d. An enlarged view of the area enclosed by the white square is shown in the inset. **(B)** Box plot of the number of ectopic xylem vessel cells per cotyledon ($n = 20$). Asterisks indicate statistically significant differences compared with the mock control (Student's *t*-test, $p < 0.0001$). **(C)** Cotyledons were cut from seven-day-old of wild-type, *knat7-1*, *blh6-1*, *knat7 blh6*, and *ofp1 ofp2 ofp3 ofp4 ofp5* (*ofp12345*) seedlings, and ectopic xylem vessel cell differentiation was induced by phytohormonal treatment (the KDB system: Tan et al., 2018) with DMSO (mock) or 5 μM TSA. **(D)** Close-up views of the area enclosed by the red square in **(C)**. **(E)** Box plot of the number of ectopic xylem vessel cells per cotyledon under the TSA treatment ($n = 13$). Different characters indicate statistically significant differences (Tukey's test, $p < 0.05$). White triangles, endogenous xylem vessels; white arrows, ectopic xylem vessel cells transdifferentiated from mesophyll cells; yellow triangles, ectopic xylem vessel cells originating from vascular cells around endogenous xylem vessels; red triangles, ectopic xylem vessel cells located near endogenous xylem vessels originating from non-vascular cells. In box plots, center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots [$n = 20$ in **(B)**, and 13 in **(E)**]. Bars, 200 μm **(A,C)** and 50 μM [insets of **(A,D)**].

methylation in xylem cell formation was reported in *Eucalyptus grandis* (Hussey et al., 2015, 2017) and in Arabidopsis (Wang et al., 2021). In the case of Arabidopsis stem development, Wang et al. (2021) demonstrated that ARABIDOPSIS HOMOLOG of TRITHORAX1 (ATX1), a H3K4-histone methyltransferase, directly regulates the H3k4me3 methylation levels of the gene loci for *SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1* (*SND1*) and *NAC SECONDARY WALL THICKENING PROMOTING FACTOR1* (*NST1*), which are critical transcriptional factors for fiber cell differentiation

(Zhong et al., 2006; Mitsuda et al., 2007). *SND1* and *NST1* are included in the sister group to VND family proteins (Kubo et al., 2005; Zhong et al., 2006; Mitsuda et al., 2007; Nakano et al., 2015). Therefore, we postulate that both histone methylation and histone acetylation are important for the modification of the NAC–MYB-based transcriptional network for xylem cell formation (Nakano et al., 2015; Ohtani and Demura, 2019). Future analysis will clarify the details of histone modification-based regulation of xylem vessel cell differentiation.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

RH, TD, and MO designed the experiments. RH performed the screening of chemicals and gene expression analysis. RH, SW, and MO performed the mutant analysis. RH, SW, TD, and MO wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was in part supported by the RIKEN Center for Sustainable Sciences, MEXT KAKENHI (JP18H05484 and JP18H05489 to MO and TD; JP20H05405 and JP21H05652 to MO), JSPS KAKENHI (JP20H03271 to MO and JP18H02466 to TD), ERATO JST (JPMJER1602 to MO), the Toray Science Foundation (No. 19-6002 to MO), the Naito Foundation (to MO), and the Asahi Glass Foundation (to MO).

ACKNOWLEDGMENTS

We thank Dr. Masatoshi Yamaguchi (Saitama University, Japan) and Dr. Yoichiro Watanabe (NAIST, Japan) for providing mutant seeds; Dr. Arata Yoneda, Dr. Ko Kato, Dr. Minoru Kubo, Dr. Tadashi Kunieda, Dr. Miyuki Nakata, Dr. Masaaki Umeda, and Dr. Satoko Yoshida (NAIST, Japan) for their fruitful discussions; and Dr. Harunori Kawabe, Ms. Seiko Kinjyo, Ms. Shizuka Nishida, Ms. Eriko Tanaka, Ms. Yuki Mitsubayasi, and Ms. Ayumi Ihara (NAIST, Japan) for their technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.825810/full#supplementary-material

REFERENCES

Ageeva-Kieferle, A., Georgii, E., Winkler, B., Ghirardo, A., Albert, A., Hüther, P., et al. (2021). Nitric oxide coordinates growth, development, and stress response via histone modification and gene expression. *Plant Physiol.* 187, 336–360. doi: 10.1093/plphys/kiab222

Avci, U., Earl, P. H., Ismail, I. O., Beers, E. P., and Haigler, C. H. (2008). Cysteine proteases XCP1 and XCP2 aid micro-autolysis within the intact central vacuole during xylogenesis in Arabidopsis roots. *Plant J.* 56, 303–315. doi: 10.1111/j.1365-313X.2008.03592.x

Bhargava, A., Ahad, A., Wang, S., Mansfield, S. D., Haughn, G. W., Douglas, C. J., et al. (2013). The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in Arabidopsis. *Planta* 237, 1199–1211. doi: 10.1007/s00425-012-1821-9

Bhargava, A., Mansfield, S. D., Hall, H. C., Douglas, C. J., and Ellis, B. E. (2010). MYB75 functions in regulation of secondary cell wall formation in the Arabidopsis inflorescence stem. *Plant Physiol.* 154, 1428–1438. doi: 10.1104/pp.110.162735

Bollhöner, B., Zhang, B., Stael, S., Denancé, N., Overmyer, K., Goffner, D., et al. (2013). Post mortem function of AtMC9 in xylem vessel elements. *N. Phytol.* 200, 498–510. doi: 10.1111/nph.12387

Bourque, S., Dutartre, A., Hammoudi, V., Blanc, S., Dahan, J., Jeandroz, S., et al. (2011). Type-2 histone deacetylases as new regulators of elicitor-induced cell death in plants. *N. Phytol.* 192, 127–139. doi: 10.1111/j.1469-8137.2011.03788.x

Brown, D. M., Zeef, L. A. H., Ellis, J., Goodacre, R., and Turner, S. R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell* 17, 2281–2295. doi: 10.1105/tpc.105.031542

Chang, S., and Pikaard, C. S. (2005). Transcript profiling in Arabidopsis reveals complex responses to global inhibition of DNA methylation and histone deacetylation. *J. Biol. Chem.* 280, 796–804. doi: 10.1074/jbc.M409053200

Comer, A. E. (1978). Pattern of cell division and wound vessel member differentiation in coleus pith explants. *Plant Physiol.* 62, 354–359. doi: 10.1104/pp.62.3.354

Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, A., et al. (2015). Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. *Plant Cell Physiol.* 56, 242–254. doi: 10.1093/pcp/pcu134

Funk, V., Kositsup, B., Zhao, C., and Beers, E. P. (2002). The Arabidopsis xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. *Plant Physiol.* 128, 84–94. doi: 10.1104/pp.010514

Furuya, T., Saito, M., Uchimura, H., Satake, A., Nosaki, S., Miyakawa, T., et al. (2021). Gene co-expression network analysis identifies BEH3 as a stabilizer of secondary vascular development in Arabidopsis. *Plant Cell* 33, 2618–2636. doi: 10.1093/plcell/koab151

Grozinger, C. M., Chao, E. D., Blackwell, H. E., Moazed, D., and Schreiber, S. L. (2001). Identification of a class of small molecule inhibitors of the sirtuin family of NAD-dependent deacetylases by phenotypic screening. *J. Biol. Chem.* 276, 38837–38843. doi: 10.1074/jbc.M106779200

Hasegawa, J., Sakamoto, Y., Nakagami, S., Aida, M., Sawa, S., and Matsunaga, S. (2016). Three-dimensional imaging of plant organs using a simple and rapid transparency technique. *Plant Cell Physiol.* 57, 462–472. doi: 10.1093/pcp/pcw027

Henmi, K., Demura, T., Tsuboi, S., Fukuda, H., Iwabuchi, M., and Ogawa, K. (2005). Change in the redox state of glutathione regulates differentiation of tracheary elements in zinnia cells and arabidopsis roots. *Plant Cell Physiol.* 46, 1757–1765. doi: 10.1093/pcp/pci198

Henmi, K., Tsuboi, S., Demura, T., Fukuda, H., Iwabuchi, M., and Ogawa, K. I. (2001). A possible role of glutathione and glutathione disulfide in tracheary element differentiation in the cultured mesophyll cells of *zinnia elegans*. *Plant Cell Physiol.* 42, 673–676. doi: 10.1093/pcp/pce072

Hilal, M., Zenoff, A. M., Ponessa, G., Moreno, H., and Massa, E. M. (1998). Saline stress alters the temporal patterns of xylem differentiation and alternative oxidase expression in developing soybean roots. *Plant Physiol.* 117, 695–701. doi: 10.1104/pp.117.2.695

Hirai, R., Higaki, T., Takenaka, Y., Sakamoto, Y., Hasegawa, J., Matsunaga, S., et al. (2019). The progression of xylem vessel cell differentiation is dependent on the activity level of VND7 in *Arabidopsis thaliana*. *Plants* 9:39. doi: 10.3390/plants9010039

Hollender, C., and Liu, Z. (2008). Histone deacetylase genes in Arabidopsis development. *J. Integr Plant Biol.* 50, 875–885. doi: 10.1111/j.1744-7909.2008.00704.x

Hussey, S. G., Loots, M. T., van der Merwe, K., Mizrachi, E., and Myburg, A. A. (2017). Integrated analysis and transcript abundance modelling of H3K4me3 and H3K27me3 in developing secondary xylem. *Sci. Rep.* 7:3370. doi: 10.1038/s41598-017-03665-1

Hussey, S. G., Mizrachi, E., Creux, N. M., and Myburg, A. A. (2013). Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. *Front. Plant Sci.* 4:325. doi: 10.3389/fpls.2013.00325

Hussey, S. G., Mizrachi, E., Groover, A., Berger, D. K., and Myburg, A. A. (2015). Genome-wide mapping of histone H3 lysine 4 trimethylation in *Eucalyptus grandis* developing xylem. *BMC Plant Biol.* 15:117. doi: 10.1186/s12870-015-0499-0

Jacobs, W. P. (1952). The role of auxin in differentiation of xylem around a wound. *Am. J. Bot.* 39:301. doi: 10.1002/j.1537-2197.1952.tb14277.x

Kamon, E., and Ohtani, M. (2021). Xylem vessel cell differentiation: a best model for new integrative cell biology? *Curr. Opin. Plant. Biol.* 64:102135. doi: 10.1016/j.pbi.2021.102135

Kawabe, H., Ohtani, M., Kurata, T., Sakamoto, T., and Demura, T. (2018). Protein S-nitrosylation regulates xylem vessel cell differentiation in Arabidopsis. *Plant Cell Physiol.* 59, 17–29. doi: 10.1093/pcp/pcx151

Kim, W. C., Ko, J. H., and Han, K. H. (2012). Identification of a cis-acting regulatory motif recognized by MYB46, a master regulator of secondary wall biosynthesis. *Plant Mol. Biol.* 78, 489–501. doi: 10.1007/s11103-012-9880-7

Ko, J. H., Kim, W. C., and Han, K. H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J.* 60, 649–665. doi: 10.1111/j.1365-313X.2009.03989.x

Ko, J. H., Kim, W. C., Kim, J. Y., Ahn, S. J., and Han, K. H. (2012). MYB46-mediated transcriptional regulation of secondary cell wall biosynthesis. *Mol. Plant* 5, 961–963. doi: 10.1093/mp/sss076

Kondo, Y., Nurani, A. M., Saito, C., Ichihashi, Y., Saito, M., Yamazaki, K., et al. (2016). Vascular Cell Induction Culture System Using Arabidopsis Leaves (VISUAL) reveals the sequential differentiation of sieve element-like cells. *Plant Cell* 28, 1250–1262. doi: 10.1105/tpc.16.00027

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305

Li, E., Bhargava, A., Qiang, W., Friedmann, M. C., Forneris, N., Savidge, R. A., et al. (2012). The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in Populus. *N. Phytol.* 194, 102–115. doi: 10.1111/j.1469-8137.2011.04016.x

Li, E., Wang, S., Liu, Y., Chen, J. G., and Douglas, C. J. (2011). OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. *Plant J.* 67, 328–341. doi: 10.1111/j.1365-313X.2011.04595.x

Liu, X., Wei, W., Zhu, W., Su, L., Xiong, Z., Zhou, M., et al. (2017). Histone deacetylase AtSRT1 links metabolic flux and stress response in Arabidopsis. *Mol. Plant* 10, 1510–1522. doi: 10.1016/j.molp.2017.10.010

Liu, Y., and Douglas, C. J. (2015). A role for OVATE FAMILY PROTEIN1 (OFP1) and OFP4 in a BLH6-KNAT7 multi-protein complex regulating secondary cell wall formation in *Arabidopsis thaliana*. *Plant Signal. Behav.* 10:e1033126. doi: 10.1080/15592324.2015.1033126

Liu, Y., You, S., Taylor-Teeples, M., Li, W. L., Schuetz, M., Brady, S. M., et al. (2014). BEL1-LIKE HOMEODOMAIN6 and KNOTTED ARABIDOPSIS THALIANA7 interact and regulate secondary cell wall formation via repression of REVOLUTA. *Plant Cell* 26, 4843–4861. doi: 10.1105/tpc.114.128322

McCahill, I. W., and Hazen, S. P. (2019). Regulation of cell wall thickening by a Medley of mechanisms. *Trends Plant Sci.* 24, 853–866. doi: 10.1016/j.tplants.2019.05.012

Mengel, A., Ageeva, A., Georgii, E., Bernhardt, J., Wu, K., Durner, J., et al. (2017). Nitric oxide modulates histone acetylation at stress genes by inhibition of histone deacetylases. *Plant Physiol.* 173, 1434–1452. doi: 10.1104/pp.16.01734

Meyer, A. J., and Hell, R. (2005). Glutathione homeostasis and redox-regulation by sulfhydryl groups. *Photosynth Res.* 86, 435–457. doi: 10.1007/s11120-005-8425-1

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *Plant Cell* 19, 270–280. doi: 10.1105/tpc.106.047043

Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N. A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6:288. doi: 10.3389/fpls.2015.00288

Ning, Y. Q., Chen, Q., Lin, R. N., Li, Y. Q., Li, L., Chen, S., et al. (2019). The HDA19 histone deacetylase complex is involved in the regulation of flowering time in a photoperiod-dependent manner. *Plant J.* 98, 448–464. doi: 10.1111/tpj.14229

Ohashi-Ito, K., and Fukuda, H. (2010). Transcriptional regulation of vascular cell fates. *Curr. Opin. Plant Biol.* 13, 670–676. doi: 10.1016/j.pbi.2010.08.011

Ohashi-Ito, K., Iwamoto, K., and Fukuda, H. (2018). LOB DOMAIN-CONTAINING PROTEIN 15 positively regulates expression of VND7, a master regulator of tracheary elements. *Plant Cell Physiol.* 59, 989–996. doi: 10.1093/pcp/pcy036

Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 directly regulates the genes that govern programmed cell death and secondary wall formation during xylem differentiation. *Plant Cell* 22, 3461–3473. doi: 10.1105/tpc.110.075036

Ohtani, M., and Demura, T. (2019). The quest for transcriptional hubs of lignin biosynthesis: beyond the NAC-MYB-gene regulatory network model. *Curr. Opin. Biotechnol.* 56, 82–87. doi: 10.1016/j.copbio.2018.10.002

Ohtani, M., Kawabe, H., and Demura, T. (2018). Evidence that thiol-based redox state is critical for xylem vessel cell differentiation. *Plant Signal. Behav.* 13:e1428512. doi: 10.1080/15592324.2018.1428512

Peña, M. J., Zhong, R., Zhou, G. K., Richardson, E. A., O'Neil, M. A., Darvill, A. G., et al. (2007). Arabidopsis irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19, 549–563. doi: 10.1105/tpc.106.049320

Preston, J., Wheeler, J., Heazlewood, J., Li, S. F., and Parish, R.W. (2004). AtMYB32 is required for normal pollen development in arabidopsis thaliana. *Plant J.* 40, 979–95. doi: 10.1111/j.1365-313X.2004.02280.x

Reusche, M., Thole, K., Janz, D., Truskina, J., Rindfleisch, S., Drübert, C., et al. (2012). Verticillium infection triggers VASCULAR-RELATED NAC DOMAIN7-dependent *de novo* xylem formation and enhances drought tolerance in Arabidopsis. *Plant Cell* 24, 3823–3837. doi: 10.1105/tpc.112.103374

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

Soyano, T., Thitamadee, S., Machida, Y., and Chua, N. H. (2008). ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30 and ASL20/LBD18 regulate tracheary element differentiation in Arabidopsis. *Plant Cell.* 20, 3359–3373. doi: 10.1105/tpc.108.061796

Tan, T. T., Demura, T., and Ohtani, M. (2019). Creating vessel elements *in vitro*: towards a comprehensive understanding of the molecular basis of xylem vessel element differentiation. *Plant Biotechnol.* 36, 1–6. doi: 10.5511/plantbiotechnology.18.1119b

Tan, T. T., Endo, H., Sano, R., Kurata, T., Yamaguchi, M., Ohtani, M., et al. (2018). Transcription factors VND1-VND3 contribute to cotyledon xylem vessel formation. *Plant Physiol.* 176, 773–789. doi: 10.1104/pp.17.00461

Taylor-Teeples, M., Lin, L., de Lucas, M., Turco, G., Toal, T. W., Gaudinier, A., et al. (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* 517, 571–575. doi: 10.1038/nature14099

Thévenaz, P., and Unser, M. (2007). User-friendly semiautomated assembly of accurate image mosaics in microscopy. *Microsc. Res Tech*. 70, 135–146. doi: 10.1002/jemt.20393

Turco, G. M., Rodriguez-Medina, J., Siebert, S., Han, D., Valderrama-Gómez, M. Á., Vahldick, H., et al. (2019). Molecular mechanisms driving switch behavior in xylem cell differentiation. *Cell Rep*. 28, 342.e4–351.e4. doi: 10.1016/j.celrep.2019.06.041

Turner, S., Gallois, P., and Brown, D. (2007). Tracheary element differentiation. *Annu. Rev. Plant Biol*. 58, 407–433. doi: 10.1146/annurev.arplant.57.032905.105236

Ueda, M., Matsui, A., Tanaka, M., Nakamura, T., Abe, T., Sako, K., et al. (2017). The distinct roles of Class I and II RPD3-like histone deacetylases in salinity stress response. *Plant Physiol*. 175, 1760–1773. doi: 10.1104/pp.17.01332

Wang, S., Chang, Y., Guo, J., and Chen, J. G. (2007). Arabidopsis ovate family protein 1 is a transcriptional repressor that suppresses cell elongation. *Plant J*. 50, 858–872. doi: 10.1111/j.1365-313X.2007.03096.x

Wang, S., Yamaguchi, M., Grienenberger, E., Martone, P. T., Samuels, A. L., and Mansfield, S. D. (2020). The Class II KNOX genes KNAT3 and KNAT7 work cooperatively to influence deposition of secondary cell walls that provide mechanical support to Arabidopsis stems. *Plant J*. 101, 293–309. doi: 10.1111/tpj.14541

Wang, X., Wang, D., Xu, W., Kong, L., Ye, X., Zhuang, Q., et al. (2021). Histone methyltransferase ATX1 dynamically regulates fiber secondary cell wall biosynthesis in Arabidopsis inflorescence stem. *Nucleic Acids Res*. 49, 190–205. doi: 10.1093/nar/gkaa1191

Yamaguchi, M., Goué, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J. C., et al. (2010a). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 effectively induce transdifferentiation into xylem vessel elements under control of an induction system. *Plant Physiol*. 153, 906–914. doi: 10.1104/pp.110.154013

Yamaguchi, M., Kubo, M., Fukuda, H., and Demura, T. (2008). VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in Arabidopsis roots and shoots. *Plant J*. 55, 652–664. doi: 10.1111/j.1365-313X.2008.03533.x

Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., and Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of broad range of genes for xylem vessel formation. *Plant J*. 66, 579–590. doi: 10.1111/j.1365-313X.2011.04514.x

Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010b). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048

Yoshida, M., and Horinouchi, S. (1999). Trichostatin and leptomycin. Inhibition of histone deacetylation and signal-dependent nuclear export. *Ann. N. Y. Acad. Sci*. 886, 23–36. doi: 10.1111/j.1749-6632.1999.tb09397.x

Zhang, F., Wang, L., Ko, E. E., Shao, K., and Qiao, H. (2018). Histone deacetylases SRT1 and SRT2 Interact with ENAP1 to mediate ethylene-induced transcriptional repression. *Plant Cell* 30, 153–166. doi: 10.1105/tpc.17.00671

Zhang, Q., Luo, F., Zhong, Y., He, J., and Li, L. (2020). Modulation of NAC transcription factor NST1 activity by XYLEM NAC DOMAIN1 regulates secondary cell wall formation in Arabidopsis. *J. Exp. Bot*. 71, 1449–1458. doi: 10.1093/jxb/erz513

Zhang, Y., Yin, B., Zhang, J., Cheng, Z., Liu, Y., Wang, B., et al. (2019). Histone deacetylase HDT1 is involved in stem vascular development in Arabidopsis. *Int. J. Mol. Sci*. 20:3452. doi: 10.3390/ijms20143452

Zhao, C., Avci, U., Grant, E. H., Haigler, C. H., and Beers, E. P. (2007). XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J*. 53, 425–436. doi: 10.1111/j.1365-313X.2007.03350.x

Zhao, Y., Dai, X., Blackwell, H. E., Schreiber, S. L., and Chory, J. (2003). SIR1, an upstream component in auxin signaling identified by chemical genetics. *Science* 301, 1107–1110. doi: 10.1126/science.1084161

Zhong, R., Demura, T., and Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell* 18, 3158–3170. doi: 10.1105/tpc.106.047399

Zhong, R., Lee, C., and Ye, Z. H. (2010). Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. *Mol. Plant*. 3, 1087–1103. doi: 10.1093/mp/ssq062

Zhong, R., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suit of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol*. 53, 368–380. doi: 10.1093/pcp/pcr185

Zhou, J., Lee, C., Zhong, R., and Ye, Z. H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. *Plant Cell* 21, 248–266. doi: 10.1105/tpc.108.063321

Zhou, J., Zhong, R., and Ye, Z. H. (2014). Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS ONE* 9:e105726. doi: 10.1371/journal.pone.0105726

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

REVIEW
published: 03 February 2022
doi: 10.3389/fpls.2021.826871

Epigenetic Regulation of Megaspore Mother Cell Formation

*Ting Jiang and Binglian Zheng**

State Key Laboratory of Genetic Engineering, Ministry of Education Key Laboratory of Biodiversity Sciences and Ecological Engineering, Collaborative Innovation Center of Genetics and Development, School of Life Sciences, Fudan University, Shanghai, China

In flowering plants, the female gametophyte (FG) initiates from the formation of the megaspore mother cell (MMC). Among a pool of the somatic cells in the ovule primordium, only one hypodermal cell undergoes a transition of cell fate to become the MMC. Subsequently, the MMC undergoes a series of meiosis and mitosis to form the mature FG harboring seven cells with eight nuclei. Although *SPL/NZZ*, the core transcription factor for MMC formation, was identified several decades ago, which and why only one somatic cell is chosen as the MMC have long remained mysterious. A growing body of evidence reveal that MMC formation is associated with epigenetic regulation at multiple layers, including dynamic distribution of histone variants and histone modifications, small RNAs, and DNA methylation. In this review, we summarize the progress of epigenetic regulation in the MMC formation, emphasizing the roles of chromosome condensation, histone variants, histone methylation, small RNAs, and DNA methylation.

Keywords: MMC, epigenetic regulation, small RNA, DNA methylation, ovule development epigenetic regulation in MMC

OPEN ACCESS

Edited by:
Mingli Xu,
University of South Carolina, United States

Reviewed by:
Yun Ju Kim,
Institute for Basic Science (IBS), South Korea
Lin Xu,
Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences (CAS), China

***Correspondence:**
Binglian Zheng
zhengbl@fudan.edu.cn

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

Received: *01 December 2021*
Accepted: *31 December 2021*
Published: *03 February 2022*

Citation:
Jiang T and Zheng B (2022) Epigenetic Regulation of Megaspore Mother Cell Formation. Front. Plant Sci. 12:826871. doi: 10.3389/fpls.2021.826871

INTRODUCTION

Different from that in animals, the germline cells are not specialized during embryo development in plants. Instead, when plants grow from vegetative growth to reproductive growth, several specific somatic cells undergo cell fate transition to become the germline cells. In flowering plants, the male and female gametophytes (FGs) develop within the anther and the ovule, respectively. In most angiosperms and gymnosperms, only one somatic cell in the nucellus region of the ovule changes its cell identity and later becomes the megaspore mother cell (MMC). MMC undergoes two meiotic divisions to give rise to four megaspores. Then, only one megaspore near the chalaza becomes the functional megaspore (FM), while the other three cells undergo programmed cell death. Subsequently, the FM undergoes three mitotic nuclear divisions, finally resulting in the formation of a mature FG, so called embryo sac (Grossniklaus and Schneitz, 1998). As the first step of FG development, cell fate transition of MMC is of great importance.

In Arabidopsis, the pre-meiosis ovule can be divided into three parts along a proximal–distal axis, including nucellus, chalaza, and funiculus (Schneitz et al., 1997). The cells in the nucellus region can be further divided into two layers, the epidermal layer (L1) and the subepidermal layer (L2). In general, the archespore that arises from the most distal cell in L2 changes its cell fate to develop into MMC. Subsequently, the MMC becomes recognizable as a single, large, and elongated subepidermal cell, which is centrally positioned within the nucellus and displays a prominent

nucleus and nucleolus (Schneitz et al., 1997; Hernandez-Lagana et al., 2021; **Figure 1**). However, the mechanism of MMC formation remains unclear, especially, which, why, and how only one somatic cell is allowed to become MMC? In general, MMC formation is thought to be controlled by two steps: first, restricting only one cell differentiation to MMC, and second, preventing self-renewal of the designated MMC before meiosis. Here, we review major advances in the cell fate control of MMC, emphasizing the roles of epigenetic regulations, including the change of chromosome condensation status, distribution of histone variants and histone modifications, small RNA biogenesis, and DNA methylation.

KEY DEVELOPMENTAL REGULATORS OF MEGASPORE MOTHER CELL FORMATION

SPOROCYTELESS/NOZZLE (*SPL/NZZ*), a MADS-box transcriptional factor, is the first gene which was found to play a pioneer role in MMC formation, as the *spl/nzz* mutants have smaller nucellus and the archespore completely fails to undergo differentiation resulting in the complete absence of the MMC (Schiefthaler et al., 1999; Yang et al., 1999; Balasubramanian and Schneitz, 2000). In contrast, a recent study shows that ectopic expression of *SPL/NZZ* caused additional enlarged MMC-like cells in the early ovules (Mendes et al., 2020). Of note, as a pioneer transcription factor in germline formation, SPL/NZZ is also required for male gametophyte development, as microsporocyte formation was blocked in the *SPL/NZZ* mutants (Schiefthaler et al., 1999; Yang et al., 1999). The homologs of *SPL/NZZ* in tomato and rice are also essential for both male and FG development (Rojas-Gracia et al., 2017; Ren et al., 2018). SPL/NZZ uses its EAR motif to recruit co-repressor TOPLESS, to regulate sporocyte formation (Chen et al., 2014; Wei et al., 2015). Moreover, WUSCHEL (WUS), a key regulator for stem cell fate in plants, acts in concert with SPL/NZZ to contribute MMC formation (Lieber et al., 2011). Based on the observations that *SPL/NZZ* is mainly expressed at the tip of the ovule primordium (Mendes et al., 2020; Zhao et al., 2020), and WUS preferentially accumulates in the nucellar cells surrounding the MMC (Zhao et al., 2017; He et al., 2019; Mendes et al., 2020), it is thought that the roles of both SPL/NZZ and WUS in regulating MMC formation are non-cell-autonomous (**Figure 1**).

Once MMC specification is determined, the MMC undergoes meiosis to produce the four megaspores and only one of megaspores called FM develops into the mature FG via several rounds of mitoses (Grossniklaus and Schneitz, 1998). However, why the MMC is able to switch mitotic division to meiotic division? Cyclin-dependent kinase (CDK) inhibitor KIP-RELATED PROTEIN (KRP) family inhibit CDKA;1 to ensure the entry of MMC into meiosis rather than mitosis (Zhao et al., 2017; **Figure 1**). By analyzing MMC formation in the triple mutant of *KRP*, Zhao et al. (2017), shows that KRPs are essential for the restriction of the plant germline harboring only one MMC per ovule by inhibiting CDKA;1. Furthermore, CDKA;1 targets *RETINOBLASTOMA-RELATED 1* (*RBR1*), a Retinoblastoma (Rb) homolog in Arabidopsis (Ebel et al., 2004), to inhibit the designated meiocytes entering mitosis (Zhao et al., 2017). As a result, the meiocytes of the *rbr1* mutants undergo several mitotic divisions, resulting in the formation of supernumerary meiocytes that give rise to multiple MMCs per ovule (Zhao et al., 2017). Intriguingly, the expression of *WUS* expands from the surrounding somatic cells to the MMC in both *krp* and *rbr1* mutants (Zhao et al., 2017). Moreover, loss-of-function of *WUS* significantly restored the phenotype of multiple MMCs in the *rbr1* mutants (Zhao et al., 2017). However, ectopic expression of *WUS* failed to induce the entry of MMC into mitotic divisions, suggesting that RBR1 not WUS is a central hub to determine the switch of MMC differentiation (Zhao et al., 2017). In addition, RBR1 represses cell cycle regulator E2F transcription factors to regulate the cell fate of MMC, as the *e2f* mutant harbors two to three MMCs per ovule primordium (Yao et al., 2018). Altogether, these findings indicate that not only MMC specification but also MMC differentiation are tightly regulated (**Figure 1**).

DE-CONDENSED CHROMATIN AND DECREASED HETEROCHROMATIN IN THE MEGASPORE MOTHER CELL

Once a specific somatic cell is chosen to develop into the MMC, both the cell itself, the nucleus, and even the nucleolar of the MMC increase significantly in size (Schneitz et al., 1997), which mark the MMC distinguishable clearly from the surrounding somatic cells. Chromatin condensation and heterochromatin formation are usually correlated to the nucleus size (van Zanten et al., 2011; Wang et al., 2013). Using non-denaturing whole-mount DNA staining and confocal imaging, She et al. (2013) showed the MMC exhibits a 60% reduction in heterochromatin content and a decreased number of chromocenters, indicating that a quick establishment of a MMC-specific chromatin state.

Histone H1, a linker histone, establishes the compaction state of an array of nucleosomes to influence the status of chromatin condensation (Osipova et al., 1980). In Arabidopsis, H1 is encoded by three genes, *H1.1*, *H1.2*, and *H1.3* (Ascenzi and Gantt, 1997). *H1.1* and *H1.2* are significantly down-regulated in the MMC, and *H1.3* is barely detected in the ovule primordia (She et al., 2013). Moreover, H1.1 and H1.2 are *de novo* incorporated into the chromatin for condensation as meiosis occurs (She et al., 2013), suggesting that the decrease of H1 might be the consequence after a somatic cell is specialized into the MMC (**Figure 2**).

Consistent with the role of H1 in chromosome condensation, loss-of-function of *H1* causes a global decrease of heterochromatin formation and transposon silencing (Zemach et al., 2013; He et al., 2019). In plants, heterochromatin formation is usually associated with decreased active histone modifications, for example, H3K4me3, and increased inactive histone modifications, such as H3K27me3 and H3K9me2 (Bender, 2004). Immunofluorescence assays show, in contrast to those in the surrounding cells, H3K4me3 is enriched to 2.7-fold in the MMC while H3K27me1, H3K9me2, and H3K27me3 reduced in the MMC (She et al., 2013), indicating

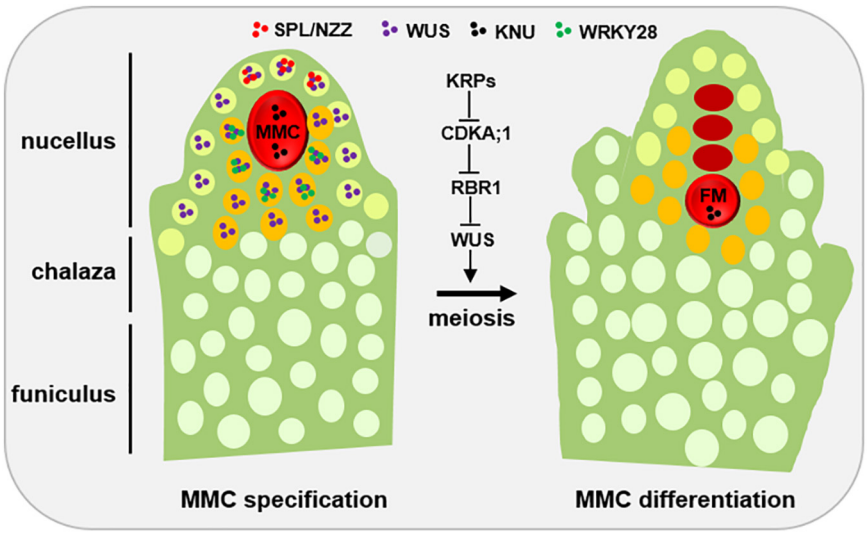

FIGURE 1 | Developmental schematic of MMC formation. In the early ovule (left), several pioneer transcription factors, mainly SPL/NZZ and WUS, play an important role in promoting MMC formation. The spatial distribution of SPL, WUS and WRKY28 away from the centered position of the nucellus region is the prerequisite of MMC specification and MMC differentiation, respectively. KNU, a widely used MMC marker. Once the MMC specification is finished, the KRP-CDKA;1-RBR1 pathway plays a key role to ensure the MMC into meiotic rather than mitotic competency by inhibiting the expansion of WUS into the MMC. L1 cell, light green; L2 cell, orange; MMC, megaspore mother cell, red; FM, functional megaspore, red. Other distal somatic cells are indicated in light white.

a permissive chromatin environment of the MMC (**Figure 2**). Correspondingly, SET DOMAIN GROUP 2 (SDG2), a writer for H3K4me3 (Berr et al., 2010; Guo et al., 2010), and LHP1, a key regulator for H3K27me3, are highly and barely expressed in the MMC, respectively (She et al., 2013). These observations indicate that with the increase of both cell size and nuclear even nucleolar size, histone modifications are actively regulated to establish a unique permission chromatin environment for the MMC.

HISTONE VARIANTS ARE ACTIVELY EXCHANGED TO CONCERT THE MEGASPORE MOTHER CELL CHROMATIN STATUS

Accompanied with chromatin condensation, histone variants often confer specific structure and functional chromatin features due to their substitutable capacity for the core canonical histone in nucleosomes in eukaryotes. Among multiple histone variants, histone H3 is encoded by *HISTONE THREE RELATED (HTR)* gene family containing 15 members in *Arabidopsis* (Okada et al., 2005). *HTR12*, a centromere-specific H3 variant CENH3 (Talbert et al., 2002), was ubiquitously expressed in the MMC (Ravi et al., 2011; She et al., 2013). By contrast, HTR8 and HTR5, two H3.3 variants that are usually associated with transcriptional competence (Ingouff et al., 2010), are specifically expressed in the MMC (She et al., 2013). HTR13, a H3.1 variant that is usually related inactive transcription activity (Jacob et al., 2014), can be gradually evicted in multiple L2 cells of the nucellus during early ovule development, but this eviction was only limited to the MMC once the identity of the MMC is designated (Hernandez-Lagana and Autran, 2020; **Figure 2**). The eviction of H3.1 in the MMC indicates that H3.1 can act as a marker to distinguish cell identity, which also happens in the root quiescent center (Otero et al., 2016). The phenomenon of multiple early L2 cells with H3.1 eviction suggests that not only one L2 cell has acquired the potential to turn into the germline cell, but finally only one can switch to the MMC by an unknown mechanism (Hernandez-Lagana and Autran, 2020).

In contrast to H3 variants, HTA11, a H2A.Z variant, is evicted from the early MMC but reincorporated later (She et al., 2013). Moreover, *WRKY28*, a transcription factor labeling the L2 cells, is activated by cytochrome P450 gene *KLU* through the chromatin remodeling complex SWR1-mediated H2A.Z deposition (Qin et al., 2014; Zhao et al., 2018). Therefore, although the mechanism by which specific chromatin hallmarks are differentially regulated in the MMC is unknown, the highly dynamic exchange among H1, H3.1, H3.3, H2A.Z, and CENH3 is consistent with a global pattern of chromatin de-condensation in the MMC, indicating that a specific chromatin reprogramming during MMC specification and differentiation (**Figure 2**).

SMALL RNA NEGATIVELY REGULATES MEGASPORE MOTHER CELL FORMATION

Based on the modes of biogenesis and action, small RNAs in plants are usually divided into three groups: microRNA (miRNA), small interfering RNA (siRNA), and trans-acting siRNA (tasiRNA) (Borges and Martienssen, 2015). In general,

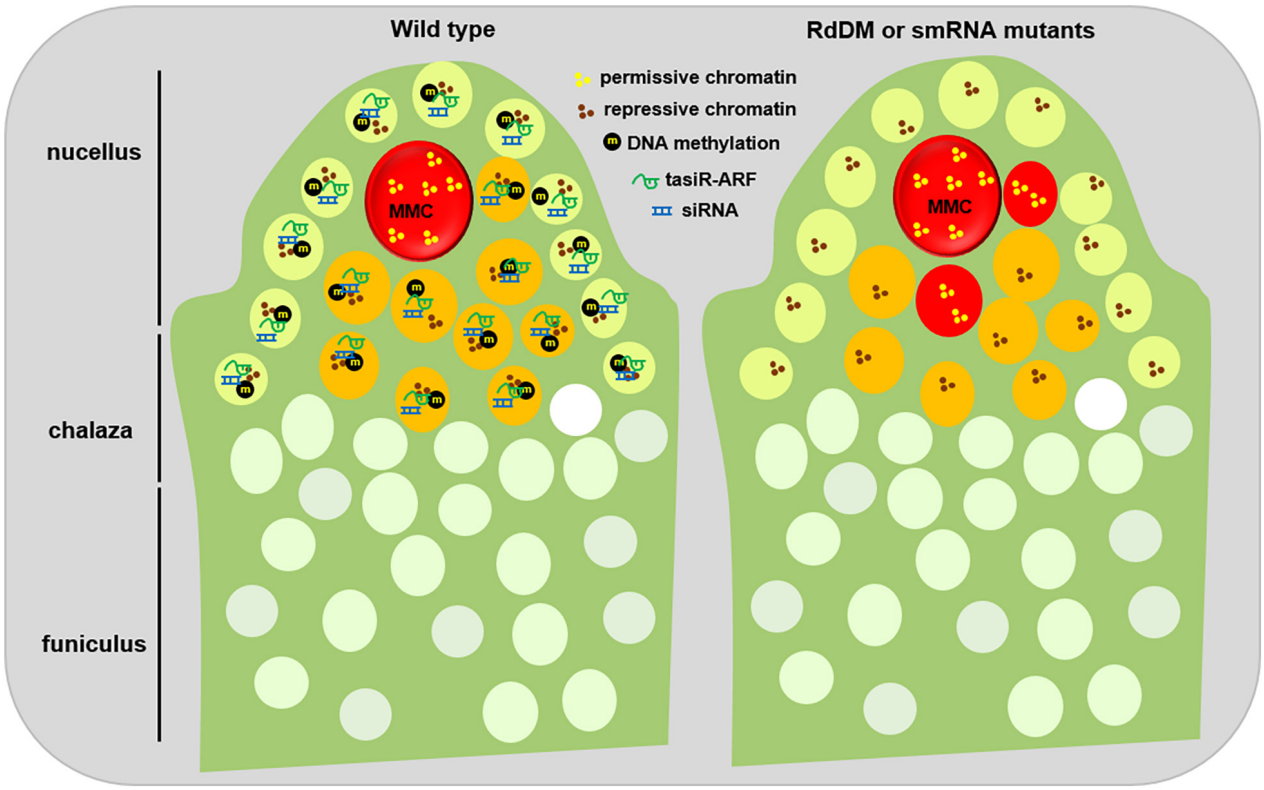

FIGURE 2 | Epigenetic regulation of MMC formation. During early ovule development of the wild type plants (left), small RNA, mainly tasiRNA and siRNA, are widely produced. On one hand, tasiRNAs impede the surrounding somatic cells to activate their potential germline identity by inhibiting ARF3 to the nucellus region; on the other hand, siRNAs reinforce the repressive chromatin status of the surrounding cells by guarding DNA methylation and/or histone modification. Meanwhile, several inactive histone modifications, such as H1, H3K9me2, and H3K27me3, facilitate the maintenance of repressive chromatin status in the surrounding cells. In contrast, the MMC shows a higher level of active histone modifications, for example, H3K4me3, thus keeping a permissive chromatin state. When the activities of small RNAs or DNA methylation are disrupted (right), the surrounding somatic cells adjacent to the MMC obtain a permissive chromatin state, and switch to the germline cell identity. This epigenetic dimorphism between the surrounding somatic cells and the MMC might ensure the establishment of a permission chromatin status in the MMC. L1 cell, light green; L2 cell, orange; MMC, megaspore mother cell, red.

MIRNA genes are transcribed into hairpin structured-precursor RNAs followed by Dicer-like 1 (DCL1)-mediated twice cleavages to produce 21–24 nt miRNAs, then miRNAs are mainly loaded onto Argonaute 1 (AGO1) for target gene inhibition with sequence complementarity (Rogers and Chen, 2013). siRNAs are mainly originated from heterochromatic regions, including transposable elements (TE) and DNA repeats. The heterochromatic regions are transcribed into double-stranded RNA precursors by Pol II-RDR6 (RNA-Dependent RNA Polymerase 6) or Pol IV-RDR2. Then, these precursors are cleaved by DCL3 to produce 21–24 nt siRNAs, which are mainly loaded onto AGO4 with the guidance of Pol V-transcribed scaffold RNAs. Lastly, the AGO4-siRNA complex recruits *de novo* DNA methyltransferase DRM2 to initiate DNA methylation for heterochromatic silencing (Matzke and Mosher, 2014). The siRNA pathway is called RdDM (RNA-directed DNA methylation). In contrast to miRNA and siRNA, tasiRNA biogenesis is initiated from specific miRNA-mediated target cleavage processes, in which non-coding *TAS* transcripts are cleaved by AGO1-miR173 or AGO7-miR390, then the cleavage products are copied into double-stranded RNAs by RDR6 with the help of SGS3 (Suppressor of Gene Silencing 3), finally these double-stranded RNAs are diced into 21 or 24 nt tasiRNA by DCL4 (Allen et al., 2005). Similar to miRNA, tasiRNAs are mainly loaded onto AGO1 to inhibit target genes.

By focusing on the function of those genes highly expressed in the FG, AGO9 was first isolated due to additional enlarged MMC-like cell formation in the *ago9* mutants (Olmedo-Monfil et al., 2010). Subsequently, further genetic analysis show that AGO4, AGO6, AGO8, other three components of the same subclass of AGO9, are all involved in MMC formation (Hernandez-Lagana et al., 2016). Consistent with the function of AGO9 in the siRNA pathway, Pol IV, RDR2, and DCL3, three key genes responsible for siRNA biogenesis, all exhibited increased incidence of additional MMC-like cells (Olmedo-Monfil et al., 2010). Of note, different ecotypes of Arabidopsis exhibit differences in the numbers of MMC, and this variation is largely correlated to the pattern differences of transcriptional regulation and localization of AGO9 in the MMC among ecotypes (Rodriguez-Leal et al., 2015). These observations demonstrate that the siRNA pathway is required to restrict the differentiation of sub-epidermal cells into the MMC in pre-meiotic ovules.

Besides those mutants in the siRNA pathway, the *rdr6*, *mir390*, *ago7*, *tas3*, mutants that affect tasiRNA biogenesis, also exhibits additional MMC-like cells per ovule (Olmedo-Monfil et al., 2010; Su et al., 2020). By screening new genes acting with RDR6 together to restrict MMC formation, TEX1, HPR1, and THOC6, several components of the THO/TREX complex, were identified as their corresponding mutants exhibit additional MMC-like cells in some pre-meiosis ovules (Su et al., 2017). The isolation of the *tho/trex* mutants is not surprising because the THO/TREX complex, similar to RDR6, is required for tasiRNA biogenesis (Jauvion et al., 2010; Yelina et al., 2010). Further evidence shows that tasiRNA inhibits the surrounding L2 cells into the MMC by restricting the expression of *Auxin Responsive Factor 3 (ARF3)* to the nucellus region (Su et al., 2017, 2020). Ectopic expression of ARF3 with *TAS3* binding site mutation in the lateral epidermal cells caused multiple MMC cells per ovule primordium (Su et al., 2020), suggesting that the inhibition of ARF3 is prerequisite for the restriction of one MMC per primordium. Moreover, these enlarged MMC-like cells of the tasiRNA mutants showed expression of *KNU*, a marker gene for MMC (Payne et al., 2004), indicating that these additional enlarged MMC-like cells have acquired the identity of MMC (Su et al., 2017). Collectively, these findings uncover the role of two small-RNA pathways in the restriction of MMC specification and differentiation (**Figure 2**).

DNA METHYLATION NEGATIVELY REGULATES MEGASPORE MOTHER CELL FORMATION

Since the siRNA pathway is required for MMC formation, and siRNA plays a role in gene silencing *via* guiding DNA methylation in plants, i.e., RdDM (Matzke and Mosher, 2014). However, little is known about DNA methylation dynamics during reproduction largely due to the technical difficulty of isolating pure and sufficient germ cells for evaluation. By developing two live imaging sensors targeting CG (MBD-Venus) and non-CG (SUVH9-Venus) methylation, respectively, Ingouff et al. (2017) showed that in contrast to the relative steady level of CG methylation during whole MMC formation, CHH methylation became undetectable in the MMC. The reduced levels of DNA methylation correlate with the de-condensed chromatin status and reduced heterochromatin formation in the MMC.

Besides siRNA biogenesis machinery (Pol IV, RDR2, and DCL3) and siRNA effectors AGO4, AGO6, and AGO9 have been involved in MMC formation, a recent finding show that the *de novo* DNA methyltransferases DRM1 and DRM2 are required for the restriction of additional MMC formation (Mendes et al., 2020), further indicating that the RdDM pathway is necessary for MMC specification and differentiation. Interestingly, *SEEDSTICK (STK)*, a MADS-box transcription factor controlling the ovule identity, binds to the CArG-box regions of *AGO9* and *RDR6* to promote their expression, and finally promoting expression of *SPL/NZZ* (Mendes et al., 2020). Moreover, in contrast to that the expression of *SPL/NZZ* is confined to the tip of early ovule/L1 layer in the wild type plants, *SPL/NZZ* ectopically expands throughout the distal nucellar primordium in the *ago9* and *drm1drm2* mutants (Mendes et al., 2020). The establishment of the STK-RdDM-SPL/NZZ relay provides direct evidence how RdDM activities is integrated by both upstream and downstream transcription factors during a specific developmental process.

Although *MET1*, a DNA methyltransferase responsible for CG methylation in Arabidopsis (Xiao et al., 2003), is ubiquitously expressed during MMC formation (Ingouff et al., 2017; Li et al., 2017), the *met1* mutant exhibits additional MMC-like cells per ovule (Li et al., 2017). Moreover, ARID1 (ARID domain-containing 1), a transcription factor that is required for heterochromatic silencing and sperm cell formation (Zheng et al., 2014), regulates MET1 reciprocally in the gamete cells, and also inhibit MMC formation (Li et al., 2017). In addition, ARID1 acts with AGO9 together to mediate siRNA movement in male gametes (Wu et al., 2021). The fact that multiple heterochromatin regulators, for example, RdDM factors, H1, MET1, and ARID1, even *TRAF Mediated Gametogenesis Progression (TRAMGaP)*, an AGO9-interacting protein (Singh et al., 2017), negatively regulate MMC specification and differentiation, indicates that heterochromatin silencing restricts the potential germline identity of the surrounding somatic cells.

EPIGENETIC REGULATION OF MEGASPORE MOTHER CELL FORMATION IN OTHER PLANTS

Although most angiosperms and gymnosperms harbor only one MMC, some plant species naturally develop more than one MMC. For example, *Trimenia moorei*, an ancient angiosperm, exhibits multiple MMCs (Bachelier and Friedman, 2011). *Gnetum*, an atypical gymnosperm, forms up to 12 MMCs and 5 of them are able to even enter meiosis (Lora et al., 2019). Why these plants develop multiple MMCs? A recent finding shows that *Utricularia gibba*, a carnivorous plant, has an unusual distribution of small RNAs and reduced global DNA methylation levels (Cervantes-Perez et al., 2021). Intriguingly, a truncated DCL3 correlates with reduced small RNA levels and DNA methylation levels, and female gametogenesis abnormalities in *U. gibba* (Cervantes-Perez et al., 2021). This finding further provides evidence that small RNA activity might be a driving force for MMC specification. Moreover, *U. gibba* might be an ideal system to investigate the evolution relationship between the RdDM pathway and MMC numbers.

A previous study ever documented the effects of natural variation of epigenetic regulators on MMC development in different ecotype of Arabidopsis (Rodriguez-Leal et al., 2015). By comparing the frequency of multiple MMCs incidence F1 hybrids of specific ecotypes, the authors show that the transcriptional patterns and protein subcellular localization of AGO9 contribute to varied MMC development among different ecotypes to an extent (Rodriguez-Leal et al., 2015). Besides Arabidopsis, several lines of evidence further show that the core genes in the small RNA pathway and DNA methylation are possibly required for MMC development. For example, loss-of-function of *dmt102* and *dmt103*, two DNA methyltransferases in maize, caused apomictic

ovule development (Garcia-Aguilar et al., 2010). AGO104, a homolog of AGO9 in maize, is also expressed in the somatic cells surrounding the MMC, and AGO104 is required for inhibiting the transition of the germline cells to the somatic cell (Singh et al., 2011). In pineapple, many genes in the RdDM pathway are highly expressed in the MMC-stage ovule (Zhao et al., 2021). Collectively, the existence of small RNA and DNA methylation-mediated gene silencing in various plant species and the expression of the corresponding genes in the ovule primordium indicate epigenetic regulation is a widely mechanism during MMC development.

CONCLUSION AND PERSPECTIVE

Considering the importance of MMC as the first female germline cell lineage in plants, to understand how this specific cell is specialized and differentiated is especially central for plant reproductive development. Classical genetic strategies have identified that several key developmental factors promote MMC specification and differentiation, such as SPL/NZZ, KRPs, RBR1, and WUS. Individual analyses of specific epigenetic regulators and epigenetic modifications show that many genes related to small RNA biogenesis and activity, DNA methylation, heterochromatin silencing, histone variants, and histone modifications, are required for MMC formation by restricting the germline identity of the surrounding somatic cells. Future work about the nature of the very beginning trigger sensed by these key factors will provide us a blueprint of the mechanism for cell fate control in plants.

Based on the differential patterns of DNA methylation, small RNA activities, and in the distribution of histone variants and histone modifications between the MMC and the surrounding somatic cells, an epigenetic dimorphism is established during MMC specification and differentiation. This dimorphism of epigenetic reprogramming might be such an above-mentioned possible trigger. Therefore, it would be very useful to create an accurate map of epigenetic dimorphism during MMC formation, if the technique difficulty of isolating high quality single cells from the early ovule primordium can be overcome in the future.

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 National Natural Science Foundation of China (32025005, 31830045, and M-0398) and the Ministry of Science and Technology of China (2021YFC2600100).

ACKNOWLEDGMENTS

We apologize for not able to refer to some related literatures due to page limit.

REFERENCES

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

Ascenzi, R., and Gantt, J. S. (1997). A drought-stress-inducible histone gene in *Arabidopsis thaliana* is a member of a distinct class of plant linker histone variants. *Plant Mol. Biol.* 34, 629–641. doi: 10.1023/a:1005886011722

Bachelier, J. B., and Friedman, W. E. (2011). Female gamete competition in an ancient angiosperm lineage. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12360–12365. doi: 10.1073/pnas.1104697108

Balasubramanian, S., and Schneitz, K. (2000). NOZZLE regulates proximal-distal pattern formation, cell proliferation and early sporogenesis during ovule development in *Arabidopsis thaliana*. *Development* 127, 4227–4238. doi: 10.1242/dev.127.19.4227

Bender, J. (2004). Chromatin-based silencing mechanisms. *Curr .Opin. Plant Biol.* 7, 521–526. doi: 10.1016/j.pbi.2004.07.003

Berr, A., McCallum, E. J., Menard, R., Meyer, D., Fuchs, J., Dong, A., et al. (2010). Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. *Plant Cell* 22, 3232–3248. doi: 10.1105/tpc.110.079962

Borges, F., and Martienssen, R. A. (2015). The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell Biol.* 16, 727–741. doi: 10.1038/nrm4085

Cervantes-Perez, S. A., Yong-Villalobos, L., Florez-Zapata, N. M. V., Oropeza-Aburto, A., Rico-Resendiz, F., Amasende-Morales, I., et al. (2021). Atypical DNA methylation, sRNA-size distribution, and female gametogenesis in *Utricularia gibba*. *Sci. Rep.* 11:15725. doi: 10.1038/s41598-021-95054-y

Chen, G. H., Sun, J. Y., Liu, M., Liu, J., and Yang, W. C. (2014). SPOROCYTELESS is a novel embryophyte-specific transcription repressor that interacts with TPL and TCP proteins in *Arabidopsis*. *J. Genet. Genom.* 41, 617–625. doi: 10.1016/j.jgg.2014.08.009

Ebel, C., Mariconti, L., and Gruissem, W. (2004). Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature* 429, 776–780. doi: 10.1038/nature02637

Garcia-Aguilar, M., Michaud, C., Leblanc, O., and Grimanelli, D. (2010). Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. *Plant Cell* 22, 3249–3267. doi: 10.1105/tpc.109.072181

Grossniklaus, U., and Schneitz, K. (1998). The molecular and genetic basis of ovule and megagametophyte development. *Semin. Cell Dev. Biol.* 9, 227–238. doi: 10.1006/scdb.1997.0214

Guo, L., Yu, Y., Law, J. A., and Zhang, X. (2010). SET DOMAIN GROUP2 is the major histone H3 lysine 4 trimethyltransferase in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18557–18562. doi: 10.1073/pnas.1010478107

He, S., Vickers, M., Zhang, J., and Feng, X. (2019). Natural depletion of histone H1 in sex cells causes DNA demethylation, heterochromatin decondensation and transposon activation. *Elife* 8:42530. doi: 10.7554/eLife.42530

Hernandez-Lagana, E., and Autran, D. (2020). H3.1 eviction marks female germline precursors in *Arabidopsis*. *Plants* 9:322. doi: 10.3390/plants9101322

Hernandez-Lagana, E., Mosca, G., Mendocilla-Sato, E., Pires, N., Frey, A., Giraldo-Fonseca, A., et al. (2021). Organ geometry channels reproductive cell fate in the *Arabidopsis* ovule primordium. *Elife* 10:66031. doi: 10.7554/eLife.66031

Hernandez-Lagana, E., Rodriguez-Leal, D., Lua, J., and Vielle-Calzada, J. P. (2016). A multigenic network of ARGONAUTE4 clade members controls early megaspore formation in *Arabidopsis*. *Genetics* 204, 1045–1056. doi: 10.1534/genetics.116.188151

Ingouff, M., Rademacher, S., Holec, S., Soljic, L., Xin, N., Readshaw, A., et al. (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic

reprogramming in *Arabidopsis*. *Curr. Biol.* 20, 2137–2143. doi: 10.1016/j.cub.2010.11.012

Ingouff, M., Selles, B., Michaud, C., Vu, T. M., Berger, F., Schorn, A. J., et al. (2017). Live-cell analysis of DNA methylation during sexual reproduction in *Arabidopsis* reveals context and sex-specific dynamics controlled by noncanonical RdDM. *Genes Dev.* 31, 72–83. doi: 10.1101/gad.289397.116

Jacob, Y., Bergamin, E., Donoghue, M. T., Mongeon, V., LeBlanc, C., Voigt, P., et al. (2014). Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* 343, 1249–1253. doi: 10.1126/science.1248357

Jauvion, V., Elmayan, T., and Vaucheret, H. (2010). The conserved RNA trafficking proteins HPR1 and TEX1 are involved in the production of endogenous and exogenous small interfering RNA in *Arabidopsis*. *Plant Cell* 22, 2697–2709. doi: 10.1105/tpc.110.076638

Li, L., Wu, W., Zhao, Y., and Zheng, B. (2017). A reciprocal inhibition between ARID1 and MET1 in male and female gametes in *Arabidopsis*. *J. Integr. Plant Biol.* 59, 657–668. doi: 10.1111/jipb.12573

Lieber, D., Lora, J., Schrempp, S., Lenhard, M., and Laux, T. (2011). *Arabidopsis* WIH1 and WIH2 genes act in the transition from somatic to reproductive cell fate. *Curr. Biol.* 21, 1009–1017. doi: 10.1016/j.cub.2011.05.015

Lora, J., Yang, X., and Tucker, M. R. (2019). Establishing a framework for female germline initiation in the plant ovule. *J. Exp. Bot.* 70, 2937–2949. doi: 10.1093/jxb/erz212

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15, 394–408. doi: 10.1038/nrg3683

Mendes, M. A., Petrella, R., Cucinotta, M., Vignati, E., Gatti, S., Pinto, S. C., et al. (2020). The RNA-dependent DNA methylation pathway is required to restrict SPOROCYTELESS/NOZZLE expression to specify a single female germ cell precursor in *Arabidopsis*. *Development* 147:194274. doi: 10.1242/dev.194274

Okada, T., Endo, M., Singh, M. B., and Bhalla, P. L. (2005). Analysis of the histone H3 gene family in *Arabidopsis* and identification of the male-gamete-specific variant AtMGH3. *Plant J.* 44, 557–568. doi: 10.1111/j.1365-313X.2005.02554.x

Olmedo-Monfil, V., Duran-Figueroa, N., Arteaga-Vazquez, M., Demesa-Arevalo, E., Autran, D., Grimanelli, D., et al. (2010). Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464, 628–632. doi: 10.1038/nature08828

Osipova, T. N., Pospelov, V. A., Svetlikova, S. B., and Vorob'ev, V. I. (1980). The role of histone H1 in compaction of nucleosomes. sedimentation behaviour of oligonucleosomes in solution. *Eur. J. Biochem.* 113, 183–188. doi: 10.1111/j.1432-1033.1980.tb06153.x

Otero, S., Desvoyes, B., Peiro, R., and Gutierrez, C. (2016). Histone H3 dynamics reveal domains with distinct proliferation potential in the *Arabidopsis* root. *Plant Cell* 28, 1361–1371. doi: 10.1105/tpc.15.01003

Payne, T., Johnson, S. D., and Koltunow, A. M. (2004). KNUCKLES (KNU) encodes a C2H2 zinc-finger protein that regulates development of basal pattern elements of the *Arabidopsis* gynoecium. *Development* 131, 3737–3749. doi: 10.1242/dev.01216

Qin, Y., Zhao, L., Skaggs, M. I., Andreuzza, S., Tsukamoto, T., Panoli, A., et al. (2014). ACTIN-RELATED PROTEIN6 regulates female meiosis by modulating meiotic gene expression in *Arabidopsis*. *Plant Cell* 26, 1612–1628. doi: 10.1105/tpc.113.120576

Ravi, M., Shibata, F., Ramahi, J. S., Nagaki, K., Chen, C., Murata, M., et al. (2011). Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*. *PLoS Genet* 7:e1002121. doi: 10.1371/journal.pgen.1002121

Ren, L., Tang, D., Zhao, T., Zhang, F., Liu, C., Xue, Z., et al. (2018). OsSPL regulates meiotic fate acquisition in rice. *New Phytol.* 218, 789–803. doi: 10.1111/nph.15017

Rodriguez-Leal, D., Leon-Martinez, G., Abad-Vivero, U., and Vielle-Calzada, J. P. (2015). Natural variation in epigenetic pathways affects the specification of female gamete precursors in *Arabidopsis*. *Plant Cell* 27, 1034–1045. doi: 10.1105/tpc.114.133009

Rogers, K., and Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* 25, 2383–2399. doi: 10.1105/tpc.113.113159

Rojas-Gracia, P., Roque, E., Medina, M., Rochina, M., Hamza, R., Angarita-Diaz, M. P., et al. (2017). The parthenocarpic hydra mutant reveals a new function for a SPOROCYTELESS-like gene in the control of fruit set in tomato. *New Phytol.* 214, 1198–1212. doi: 10.1111/nph.14433

Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E., and Schneitz, K. (1999). Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 96, 11664–11669. doi: 10.1073/pnas.96.20.11664

Schneitz, K., Hulskamp, M., Kopczak, S. D., and Pruitt, R. E. (1997). Dissection of sexual organ ontogenesis: a genetic analysis of ovule development in *Arabidopsis thaliana*. *Development* 124, 1367–1376. doi: 10.1242/dev.124.7.1367

She, W., Grimanelli, D., Rutowicz, K., Whitehead, M. W., Puzio, M., Kotlinski, M., et al. (2013). Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. *Development* 140, 4008–4019. doi: 10.1242/dev.095034

Singh, M., Goel, S., Meeley, R. B., Dantec, C., Parrinello, H., Michaud, C., et al. (2011). Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. *Plant Cell* 23, 443–458. doi: 10.1105/tpc.110.079020

Singh, S. K., Kumar, V., Srinivasan, R., Ahuja, P. S., Bhat, S. R., and Sreenivasulu, Y. (2017). The TRAF mediated gametogenesis progression (TRAMGaP) gene is required for megaspore mother cell specification and gametophyte development. *Plant Physiol.* 175, 1220–1237. doi: 10.1104/pp.17.00275

Su, Z., Wang, N., Hou, Z., Li, B., Li, D., Liu, Y., et al. (2020). Regulation of female germline specification via small RNA mobility in *Arabidopsis*. *Plant Cell* 32, 2842–2854. doi: 10.1105/tpc.20.00126

Su, Z., Zhao, L., Zhao, Y., Li, S., Won, S., Cai, H., et al. (2017). The THO complex non-cell-autonomously represses female germline specification through the TAS3-ARF3 module. *Curr Biol* 27, 1597–1609.e1592. doi: 10.1016/j.cub.2017.05.021

Talbert, P. B., Masuelli, R., Tyagi, A. P., Comai, L., and Henikoff, S. (2002). Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. *Plant Cell* 14, 1053–1066. doi: 10.1105/tpc.010425

van Zanten, M., Koini, M. A., Geyer, R., Liu, Y., Brambilla, V., Bartels, D., et al. (2011). Seed maturation in *Arabidopsis thaliana* is characterized by nuclear size reduction and increased chromatin condensation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20219–20224. doi: 10.1073/pnas.1117726108

Wang, H., Dittmer, T. A., and Richards, E. J. (2013). *Arabidopsis* CROWDED NUCLEI (CRWN) proteins are required for nuclear size control and heterochromatin organization. *BMC Plant Biol.* 13:200. doi: 10.1186/1471-2229-13-200

Wei, B., Zhang, J., Pang, C., Yu, H., Guo, D., Jiang, H., et al. (2015). The molecular mechanism of sporocyteless/nozzle in controlling *Arabidopsis* ovule development. *Cell Res.* 25, 121–134. doi: 10.1038/cr.2014.145

Wu, W., Li, L., Zhao, Y., Zhao, Y., Jiang, T., McCormick, S., et al. (2021). Heterochromatic silencing is reinforced by ARID1-mediated small RNA movement in *Arabidopsis* pollen. *New Phytol.* 229, 3269–3280. doi: 10.1111/nph.16871

Xiao, W., Gehring, M., Choi, Y., Margossian, L., Pu, H., Harada, J. J., et al. (2003). Imprinting of the MEA polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Dev. Cell* 5, 891–901. doi: 10.1016/s1534-5807(03)00361-7

Yang, W. C., Ye, D., Xu, J., and Sundaresan, V. (1999). The SPOROCYTELESS gene of *Arabidopsis* is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes Dev.* 13, 2108–2117. doi: 10.1101/gad.13.16.2108

Yao, X., Yang, H., Zhu, Y., Xue, J., Wang, T., Song, T., et al. (2018). The canonical E2Fs are required for germline development in *Arabidopsis*. *Front. Plant Sci.* 9:638. doi: 10.3389/fpls.2018.00638

Yelina, N. E., Smith, L. M., Jones, A. M., Patel, K., Kelly, K. A., and Baulcombe, D. C. (2010). Putative *Arabidopsis* THO/TREX mRNA export complex is involved in transgene and endogenous siRNA biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13948–13953. doi: 10.1073/pnas.0911341107

Zemach, A., Kim, M. Y., Hsieh, P. H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., et al. (2013). The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* 153, 193–205. doi: 10.1016/j.cell.2013.02.033

Zhao, H., Guo, M., Yan, M., Cheng, H., Liu, Y., She, Z., et al. (2020). Comparative expression profiling reveals genes involved in megasporogenesis. *Plant Physiol.* 182, 2006–2024. doi: 10.1104/pp.19.01254

Zhao, L., Cai, H., Su, Z., Wang, L., Huang, X., Zhang, M., et al. (2018). KLU suppresses megasporocyte cell fate through SWR1-mediated activation of WRKY28 expression in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 115, E526–E535. doi: 10.1073/pnas.1716054115

Zhao, L., Liu, L., Liu, Y., Dou, X., Cai, H., Aslam, M., et al. (2021). Characterization of germline development and identification of genes associated with germline specification in pineapple. *Hortic Res.* 8:239. doi: 10.1038/s41438-021-00669-x

Zhao, X., Bramsiepe, J., Van Durme, M., Komaki, S., Prusicki, M. A., Maruyama, D., et al. (2017). RETINOBLASTOMA RELATED1 mediates germline entry in *Arabidopsis*. *Science* 356:aaf6532. doi: 10.1126/science.aaf6532

Zheng, B., He, H., Zheng, Y., Wu, W., and McCormick, S. (2014). An ARID domain-containing protein within nuclear bodies is required for sperm cell formation in *Arabidopsis thaliana*. *PLoS Genet* 10:e1004421. doi: 10.1371/journal.pgen.1004421

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

MINI REVIEW
published: 10 February 2022
doi: 10.3389/fpls.2022.826473

Epigenetic Regulation of Heat Stress in Plant Male Reproduction

Shikha Malik† *and Dazhong Zhao**

Department of Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI, United States

In flowering plants, male reproductive development is highly susceptible to heat stress. In this mini-review, we summarized different anomalies in tapetum, microspores, and pollen grains during anther development under heat stress. We then discussed how epigenetic control, particularly DNA methylation, is employed to cope with heat stress in male reproduction. Further understanding of epigenetic mechanisms by which plants manage heat stress during male reproduction will provide new genetic engineering and molecular breeding tools for generating heat-resistant crops.

Keywords: heat stress, male reproduction, anther, tapetum, and pollen development, epigenetic regulation, DNA methylation

OPEN ACCESS

Edited by:
Seonghoe Jang,
World Vegetable Center Korea Office, South Korea

Reviewed by:
Biao Jin,
Yangzhou University, China
Ranjith Papareddy,
University of California, Los Angeles, United States

***Correspondence:**
Dazhong Zhao
dzhao@uwm.edu

†Present address:
Shikha Malik,
Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, United States

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

Received: *30 November 2021*
Accepted: *21 January 2022*
Published: *10 February 2022*

Citation:
Malik S and Zhao D (2022) Epigenetic Regulation of Heat Stress in Plant Male Reproduction. Front. Plant Sci. 13:826473. doi: 10.3389/fpls.2022.826473

INTRODUCTION

Short- and long-term heat stress have detrimental effects on overall growth and development in plants (Kotak et al., 2007); however, reproductive organs, particularly the male reproductive organ, are more susceptible to elevated temperatures comparing with vegetative organs (Abiko et al., 2005; Sakata et al., 2010; Sato et al., 2014, 2019; Fragkostefanakis et al., 2016; Begcy et al., 2019; He et al., 2019). Heat stress leads to partial or complete male sterility, which in turn causes yield loss in crops (Smith and Zhao, 2016). Being sessile, plants employ various mechanisms to cope with heat stress. Besides the genetic control, transcriptome and genome-wide DNA methylation analyses have revealed that the epigenetic regulation plays a pivotal role in reprogramming expression of genes required for plants to manage heat stress during reproductive development. In this mini-review, we focus on discussing research in epigenetic mechanisms underlying heat stress response in male reproduction.

PLANT MALE REPRODUCTION IS HIGHLY SENSITIVE TO HEAT STRESS

Heat stress impairs anther wall cell differentiation, microsporogenesis, and pollen formation, resulting in partial or complete male sterility in various plants. Stamen is the male reproductive organ of flowering plants, comprising of an anther where pollen (the male gametophyte) develops and a filament that anchors the anther to the flower. A typical anther has four lobes (microsporangia; Goldberg et al., 1993; Zhao, 2009; Feng et al., 2013; Walbot and Egger, 2016); within each lobe, the central pollen mother cells (PMC or microsporocytes) are surrounded by four concentrically organized layers of somatic cells: the epidermis, endothecium, middle layer, and tapetum (outside to inside). PMCs give rise to pollen *via* a series of events. PMCs undergo meiosis to produce tetrads that release microspores. After two rounds of mitosis,

microspores eventually become pollen grains which contain a vegetative cell and two sperm cells (Sanders et al., 1999; **Figure 1**). The somatic anther wall cells, particularly tapetal cells (tapetum), are essential for the normal development and release of pollen. Tapetum, consisting of a monolayer or multilayers of endopolyploid cells, which is associated with successive stages of PMC, tetrads, microspores, and developing pollen as anther development progresses (Goldberg et al., 1993; Scott et al., 2004; Walbot and Egger, 2016; **Figure 1**). Early on, tapetal cells secrete enzymes required for releasing haploid microspores from tetrads (Pacini et al., 1985; Clément and Pacini, 2001; Hsieh and Huang, 2007; Ishiguro et al., 2010; Parish and Li, 2010). Later, tapetal cells provide energy and materials for pollen development and pollen coat formation (Wu et al., 1997; Wang et al., 2003; Parish and Li, 2010; Huang et al., 2017). Lack of a tapetum or an abnormal tapetum impairs microspore and pollen development, causing male sterility (Mariani et al., 1990; Zhao et al., 2002; Zhang et al., 2014). Furthermore, endothecium is necessary for anther dehiscence (Cecchetti et al., 2013; Murphy et al., 2015).

Heat stress causes male sterility and seed yield loss are mainly ascribed to aberrant tapetum and pollen development (Parish et al., 2012; De Storme and Geelen, 2014). Decreased pollen viability due to heat stress has been reported in many crops, such as common bean (Gross and Kigel, 1994; Prasad et al., 2002), rice (Endo et al., 2009), cotton (Min et al., 2014; Song et al., 2015), tomato (Pressman et al., 2002; Giorno et al., 2013), pepper (Erickson and Markhart, 2002), wheat (Saini and Aspinall, 1982; Saini et al., 1984), barley (Sakata et al., 2010), cowpea (Ahmed et al., 1992), peanut (Vara Prasad et al., 1999; Zoong Lwe et al., 2020), and flax (Cross et al., 2003; **Table 1**). In crops, such as wheat, episodes of male sterility were observed upon 3 days of treatment at 30/30°C (day/night, the same thereafter) during meiosis, and irregular tapetum degeneration is a plausible cause for pollen abortion (Saini et al., 1984). In heat-sensitive wheat varieties, elevated temperature (35/24°C) caused tapetum degradation and pollen abortion (Browne et al., 2021). Premature pollen development in common bean at 33/29°C is also a result of early tapetum degeneration (Suzuki et al., 2001). Furthermore, abnormally wavy, looped endoplasmic reticulum (ER) structures were detected in heat-stressed tapetal cells (Suzuki et al., 2001), suggesting that ER malfunction in tapetal cells might cause male sterility under heat stress (De Storme and Geelen, 2014). Heat stress results in DNA fragmentation, cytoplasmic shrinkage, and vacuolation in early tapetal cells of thermosensitive genic male-sterile (TGMS) rice, suggesting that the precocious programmed cell death (PCD) of tapetal cells during heat stress causes male sterility (Ku et al., 2003). Impaired tapetal cells by heat stress also affects callose degradation in PMCs and pollen wall formation, such as exine patterning (Suzuki et al., 2001; Parish et al., 2012; Djanaguiraman et al., 2014). Moderately high temperature (30/25°C) causes aberrant mitochondria, ER, and nuclear membranes in PMCs (Oshino et al., 2007). Moreover, abnormal meiosis occurred in PMCs in heat-stressed wheat (Omidi et al., 2014). Recently, abnormal cross-over was observed in *Arabidopsis* male meiocytes under high temperature (De Storme and Geelen, 2020). Heat stress (36–38°C) also impaired chromosome segregation and cytokinesis during male meiosis in *Arabidopsis* (Lei et al., 2020). Moreover, acute heat stress on *Arabidopsis* causes defects in male germline and sporophytic anther tissues (Hedhly et al., 2020). A recent report showed

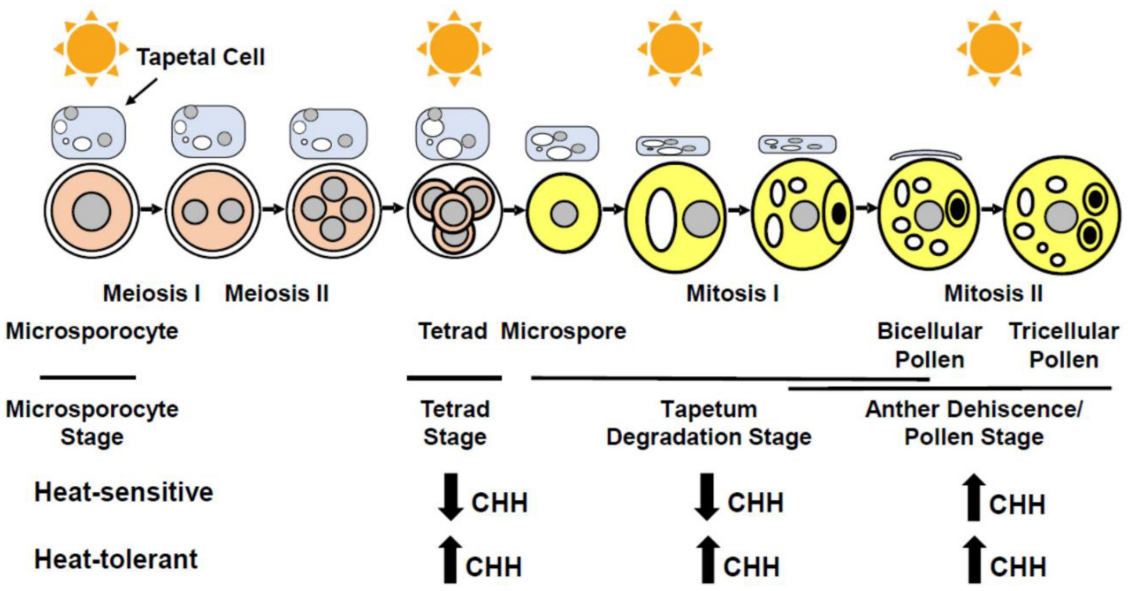

FIGURE 1 | Schematic representation of stages susceptible to heat stress (indicated by sun symbols) during male reproduction and their methylation patterns. CHH methylation is decreased under heat stress at tetrad and tapetum degradation stages in anthers of heat-sensitive plants. CHH methylation is increased under heat stress at tetrad and tapetum degradation stages in anthers of heat-tolerant plants, as well as at the anther dehiscence/pollen stage in anthers of both heat-sensitive and -tolerant plants (H in CHH representing A, T, or G).

TABLE 1 | Effects of heat stress on plant male reproduction.

Plant	Temperature	Effect	Reference
Arabidopsis (*Arabidopsis thaliana*)	30–32°C 6 to 48 h; 36–38°C; 24 h	Abnormal anther wall, male meiosis, male germline, and meiotic cytokinesis	De Storme and Geelen, 2020; Lei et al., 2020
Cotton (*Gossypium hirsutum*)	35–39°C 7 days and 40/34°C	Abnormal microspores, tapetum, and pollen grains	Min et al., 2014; Song et al., 2015; Ma et al., 2018
Barley (*Hordeum vulgare*)	30/25°C (day/night) 5 days	Abnormal pollen mother cell and tapetum	Abiko et al., 2005; Oshino et al., 2007; Sakata et al., 2010
Rice (*Oryza sativa*)	39/30°C (day/night) and 32°C	Decreased pollen viability, premature tapetum degradation in TGMS rice	Ku et al., 2003; Endo et al., 2009; Zhao et al., 2018
Maize (*Zea mays*)	35/25°C (day/night) 3 days	Decreased pollen viability	Lee et al., 2021
Tomato (*Lycopersicon esculentum*)	36/26°C (day/night) 3 days and 32/26°C	Aberrant male gametogenesis, decreased pollen grain viability	Pressman et al., 2002; Giorno et al., 2013
Wheat (*Triticum aestivum*)	>30 and 30°C for 3 days	Abnormal anthers, tapetum degradation, sporogenesis, and pollen grain viability	Saini and Aspinall, 1982; Saini et al., 1984; Omidi et al., 2014; Browne et al., 2021
Bean (*Phaseolus vulgaris*)	32.7 and 32/27°C (1 or 5 days)	Abnormal pollen grains and tapetum	Gross and Kigel, 1994; Suzuki et al., 2001
Cowpea (*Vigna unguiculata*)	33/20°C or 33/30°C (day/night)	Tapetum, tetrads disorganized, Abnormal pollen grains	Ahmed et al., 1992
Bell pepper (*Capsicum annuum*)	36°C	Deformed pollen grains	Erickson and Markhart, 2002
Brachypodium distachyon	36°C	Aborted uninucleate, vacuolated microspore, ruptured tapetal cells, Abnormal pollen grains	Harsant et al., 2013
Flax (*Linum usistatissimum*)	Increase of 3°C per hour to 40°C for 7 h, held for 2 h at 40°C	Compressed and folded pollen grains	Cross et al., 2003
Grain sorghum (*Sorghum bicolor*)	36/26°C and 38/28°C for 10 days	Reduced pollen germination	Jain et al., 2007; Djanaguiraman et al., 2014
Peanut (*Arachis hypogaea*)	28, 34, 42, and 48°C	Pollen viability	Vara Prasad et al., 1999; Zoong Lwe et al., 2020

that pollen abortion was subjected to heat stress (35/25°C) at the pre-meiotic stage in maize with downregulated *MAGO* (*MALE-ASSOCIATED ARGONAUTE-1* and *-2*) genes (Lee et al., 2021). Further studies revealed that heat stress induced MAGO hypophosphorylation which affects accumulation of 21-nt phasiRNAs and then the activity of retrotransposons in anther wall cells. Thus, the surveillance mechanism mediated by Argonaute is important for protecting male sterility under heat stress.

Anther wall cells and pollen in tomato plants upon heat stress (32/26°C) witness decreased starch and soluble sugar contents (Pressman et al., 2002). In sorghum, heat-stressed (36/26°C) microspores also showed reduced starch content and sucrose deficiency, thus reducing pollen germination (Jain et al., 2007). Moreover, an imbalance in ROS (reactive oxygen species) homeostasis in tapetal cells due to heat stress possibly causes early PCD of tapetal cells (De Storme and Geelen, 2014). In rice anthers, ROS and superoxide dismutase (SOD) are significantly increased at the male meiosis stage (Zhao et al., 2018). In barley, male sterility is possibly attributed to the hyper-phosphorylation of the serine-5 residue at the C-terminal domain of RNA Polymerase (RNA Pol) II, which alters expression of many genes during early anther development under high-temperature conditions (Abiko et al., 2005). Furthermore, auxin synthesis in *Arabidopsis* and barley anthers are reduced during high temperatures, whereas exogenous application of auxin to anthers improved pollen thermotolerance in barley (Sakata et al., 2010; Higashitani, 2013). Auxin biosynthesis genes, such as *YUCCA-YUC2* and *YUC6*, were suppressed in anthers exposed to high temperatures (33°C; Sakata et al., 2010). Heat stress generally alters expression of various genes which affect cell proliferation, photosynthesis, hormones, starch metabolism, heat shock response, and ROS production (Yang et al., 2006; Yamakawa et al., 2007; Endo et al., 2009; Frank et al., 2009; Bita et al., 2011; Mangelsen et al., 2011; Guan et al., 2013; Min et al., 2014; Song et al., 2014; Fragkostefanakis et al., 2015; González-Schain et al., 2016; Zhang et al., 2017; Zhao et al., 2018; Begcy et al., 2019; Qian et al., 2019b). Here we mainly discuss the epigenetic mechanisms by which plants respond to heat stress during male reproduction.

EPIGENETIC MODIFICATIONS DURING HEAT STRESS RESPONSE

In contrast with the molecular mechanisms underlying heat stress at the transcriptional level, epigenetic regulation during high-temperature stress is not well understood in plants (Ohama et al., 2017). Different plant organs/cells have been studied to understand the role of epigenetic modifications during heat stress. For instance, exposure of soybean root hairs and roots stripped root hairs to heat stress (40°C) caused hypomethylation of CHH (H = A, T or C; Hossain et al., 2017). Heat stress also induced hypomethylation of CG and CHG in cultured microspores of *Brassica napus* (Li et al., 2016). In maize seedlings, 325 differentially methylated genes (DMG) were

identified responding to heat stress (42°C). Interestingly, 9 DMG associated with spliceosome showed the decreased methylation level during heat stress (Qian et al., 2019a). Moreover, the *Brassica napus* heat-sensitive genotype possesses a higher level of DNA methylation than the heat-tolerant genotype during heat stress (37–45°C; Gao et al., 2014). Collectively, these findings reveal that DNA methylation is responsive to heat stress. The effect of heat stress on methylation in various plants is summarized in **Table 2**.

Genes involved in DNA methylation, histone modification, chromatin modeling, and small RNA biogenesis were studied for their roles in response to heat stress. Loss-of-function mutant of the *NUCLEAR RNA POLYMERASE D2A* (*NRPD2*) gene which encodes the second largest subunit of RNA POL IV and POL V is sensitive to heat stress (Popova et al., 2013). A RPD3-type of histone deacetylase mutant *hda6* is sensitive to heat stress. In contrast, DNA methyltransferase mutants, such as *domains rearranged methylase1* (*drm1*), *domains rearranged methylase2* (*drm2*), and *chromomethylase3*(*cmt3*), presented less pronounced response to heat stress (Popova et al., 2013). Interestingly in wild-type *Arabidopsis* plants, heat stress induced expression of the key DNA methyltransferase gene *DRM2* as well as *NUCLEAR RNA POLYMERASE D1A* (*NRPD1*) and *NUCLEAR RNA POLYMERASE D1B* (*NRPE1*) which encode the largest subunit of RNA Pol IV and RNA Pol V, respectively (Naydenov et al., 2015). Conversely, the prolonged heat exposure decreased expression of DNA methyltransferase genes *METHYLASE1* (*MET1*) and *CHROMOMETHYLASE3*(*CMT3*; Naydenov et al., 2015). The *DRM2* expression during heat stress might be regulated by RNA Pol IV and/or RNA Pol V (Naydenov et al., 2015).

DNA methylation associated with NRPD2 and histone modification mediated by HDA6 might play different roles in transcriptional reprogramming for coping with heat stress. Transcriptomic analysis of directly heat-stressed *hda6* mutants revealed a larger set of mis-regulated genes comparing with the heat-stressed *nrpd2* mutant, while after recovery from heat stress a much broader transcriptional response was detected in *nrpd2* mutants than *hda6* mutants and wild-type plants (Popova et al., 2013). In *hda6* mutants, mis-regulated genes are involved in diverse functions, such as protein processing, hormone signaling, vegetative and reproductive development, transport, and metabolism; however, GO enrichment analysis found that mis-regulated genes in *nrpd2* mutants were associated with starch catabolism, fatty acid oxidation, abiotic stress response, and auxin and cytokinin signaling pathways. A little overlap of mis-regulated gene sets between *hda6* and *nrpd2* mutants suggests that *HDA6* and *NRPD2* function differently at different stages of heat response (Popova et al., 2013). Similarly, in the heat-stressed (42°C) maize seedling, some of the key KEGG pathway enrichment involve spliceosome, RNA transport, ubiquitin-mediated proteolysis, and carbon metabolism (Qian et al., 2019a), suggesting that heat stress affects a diverse range of biological pathways which might be regulated *via* the epigenetic control.

Heat stress activates the *ONSEN* ("hot spring" in Japanese) retrotransposon and synthesis of extrachromosomal DNA copies in *Arabidopsis* seedlings (Ito et al., 2011). Heat stress triggers accumulation of *ONSEN* in mutants lacking RNA Pol IV and RDR2, which are main components in the RdDM pathway. Interestingly, the memory of heat stress (i.e., transgenerational inheritance of *ONSEN* insertion) can only occur in the progeny of mutant plants defective in siRNA biogenesis. Heat stress induced epigenetic memory associated with hypermethylation of H3K4me2 and H3K4me3 can be maintained for several days in *Arabidopsis* somatic cells (Lamke et al., 2016). Moreover, transgenerational epigenetic memory induced by heat stress is transmitted *via* HEAT SHOCK TRANSCRIPTION FACTOR A2 (HSFA2) activated H3K27me3 demethylase in *Arabidopsis* (Liu et al., 2019; Yamaguchi et al., 2021). Thus, histone modification is essential for thermotolerance memory.

TABLE 2 | Methylation patterns in plants during heat stress.

Plant	Temperature	Tissue	Methylation pattern	Reference
Soybean (*Glycine max*)	40°C	Roots	Hypomethylation CHH context	Hossain et al., 2017
Rapeseed (*Brassica napus*)	37°C for 2 h and 45°C for 3 h	Seedling	Hypermethylation in heat-sensitive variety	Gao et al., 2014
Maize (*Zea mays*)	42°C for 8 h	Seedlings	Reduced methylation of 9 differentially methylated genes	Qian et al., 2019a
Rapeseed (*Brassica napus* cv. Topas)	32°C for 6 h	Cultured Microspores	Hypomethylation CG and CHG context	Li et al., 2016
Arabidopsis (*Arabidopsis thaliana*)	42°C	Leaves	Decreased DNA methylation	Korotko et al., 2021
Cotton (*Gossypium hirsutum*)	35°C to 39/29°C to 31°C day/night for 7 days	Anthers	Hypomethylation in heat-sensitive variety	Min et al., 2014; Ma et al., 2018

EPIGENETIC REGULATION OF HEAT STRESS DURING MALE REPRODUCTION

Besides genetic regulation, the epigenetic control, particularly DNA methylation, is an important mechanism for plants to manage heat stress during male reproduction. RNA-directed DNA methylation (RdDM) in plants involves various components, such as small interfering RNAs (siRNA) and DNA methyltransferase DRM2 (Law and Jacobsen, 2010). Methylation of DNA occurs at specific sites: symmetric patterns of CpG/CpNpG and asymmetric CpNpN. In plants, methylation of asymmetric cytosine (CpNpG) is regulated by *CHROMOMETHYLASE* (*CMT*; Bartee et al., 2001).

Pollen comprises one vegetative nucleus and two sperm nuclei which maintain more stable methylation patterns than

leaves and roots (Hsieh et al., 2016). The vegetative nucleus lacks *DECREASE IN DNA METHYLATION 1* (*DDM1*), leading to reactivation of transposable elements. Reduction of DNA methylation in pollen causes transcriptional reprogramming (Slotkin et al., 2009). Cell-specific DNA methylation studies revealed that CG and CHG methylation were retained in microspores and sperm cells, whereas the CHH methylation was lost (Calarco et al., 2012). Interestingly, DNA methylation is reestablished in the vegetative cell *via* siRNA-mediated RdDM (Calarco et al., 2012). Repetitive elements were found to be active during pollen development (Slotkin et al., 2009), while heat stress can activate repetitive elements in *Arabidopsis* seedlings by epigenetic regulation (Pecinka et al., 2010). Most key genes required for DNA methylation, such as *DRM2*, *NRPD1*, and *NRPE1*, are upregulated during heat stress in *Arabidopsis* (Naydenov et al., 2015), supporting the involvement of DNA methylation in heat stress. New findings suggest that sperm cells have asymmetric mCHG, whereas vegetative nuclei and microspores possess symmetric mCHG (Borges et al., 2021). DNA methylation changes during male reproductive development were recently summarized (Papareddy and Nodine, 2021).

Transcriptome studies on heat-treated cotton anthers identified various genes involved in histone modification and DNA methylation. Under heat stress, the heat-tolerant cotton line produces normal anthers and pollen, while the heat-sensitive line is defective in anther dehiscence and fails to form viable pollen. Heat stress decreased expression of *DNA CYTOSINE-5-METHYLTRANSFERASE* (*DRM1*) and *S-ADENOSYL-L-METHIONINE-DEPENDENT METHYLTRANSFERASE* (*DRM3*) at tetrad and tapetum degradation stages in heat-sensitive cotton anthers, while their expression remains similar in heat-tolerant cotton anthers with an exception of increased expression of *DRM3* at the tetrad stage (Min et al., 2014). Similarly, expression of *NEEDED FOR RDR2-INDEPENDENT DNA METHYLATION* (*NERD*), *NUCLEAR RNA POLYMERASE D1B* (*NRPD1B*), and *S-ADENOSYL-L-HOMOCYSTEINE HYDROLASE1* (*SAHH1*), which are required for normal DNA methylation, is suppressed by heat stress in heat-sensitive cotton anthers (Min et al., 2014). During heat stress, heat-sensitive cotton anthers undergo DNA hypomethylation, while heat-tolerant cotton anthers have a high level of DNA methylation. Furthermore, pollen sterility and defects in anther dehiscence are possibly caused by hypomethylation in the heat-sensitive cotton (Ma et al., 2018). Studies on expression changes of genes associated with DNA methylation in cotton anthers under heat stress provide strong evidence that the epigenetic regulation is required for plants to cope with heat stress.

CHH methylation mediated by RdDM showed more prominent changes comparing to CG and CHG methylation, suggesting that heat stress mainly induces the RdDM activity in anthers. Most of heat-induced CHH methylations were found in promoters and downstream regions of protein-coding genes (Ma et al., 2018). Interestingly, the DNA methylation status varies with anther stages upon heat stress. At tetrad, tapetum degradation, and anther dehiscence/pollen stages, the CHH methylation level in heat-tolerant cotton anthers is increased upon heat stress; however, heat-sensitive cotton anthers depicted hypo-CHH methylation patterns at tetrad and tapetum degradation stages, while an increased CHH methylation level at the anther dehiscence/pollen stage during heat stress (**Figure 1**). Hence, heat stress may affect RdDM function in an anther stage-specific manner (Ma et al., 2018). Heat stress alters the DNA methylation level, which affects expression of genes involved in sugar metabolism and ROS generation. The abnormal concentration of sugar and ROS therefore impairs anther and pollen development. These discoveries shed light on a novel molecular mechanism by which plants ensure the success of male reproduction under high temperature, thus providing new tools for improving crops to adapt to the challenge of global warming.

Long non-coding RNA (lncRNA) is important for male fertility. In rice, an lncRNA named the long-day-specific male-fertility-associated RNA (LDMAR) is essential for pollen development under the long-day condition (Ding et al., 2012). A single nucleotide mutation in *LDMAR* increased CG methylation in the *LDMAR* promoter region, which decreased the *LDMAR* expression and thus induced PCD in anther cells. The lncRNA expression responds to stresses spatially and temporally in plants (Yu et al., 2019). Among 54 putative heat stress-induced lncRNAs, *TahlnRNA27* and *TalnRNA5* were highly upregulated by heat stress in wheat (Xin et al., 2011). Differentially expressed lncRNAs were also observed during heat stress in *Brassica rapa* (Wang et al., 2019), *Brassica juncea* (Bhatia et al., 2020), and maize (Lv et al., 2019). A recent study in *Arabidopsis* showed that 131 pollen-specific intergenic expressed loci (XLOC), which mostly encode lncRNAs, are heat stress responsive (Rutley et al., 2021). These results suggest that lncRNAs might play an important role in heat stress response during male reproduction *via* epigenetic regulation.

MicroRNAs (miRNAs) are another set of non-coding RNAs which are known to regulate gene expression at the post-transcriptional level (Bartel, 2004; Liu et al., 2010; Chen et al., 2016; Huang et al., 2016). In *Brassica rapa* seedlings, heat stress significantly decreased expression of novel miRNAs *bra*-miR1885b.3 and bra-miR5716 (Yu et al., 2011). In barley, heat stress induced expression of miR160a, 166a, 167h, and 5175a, while expression levels of their target genes, such as *AUXIN RESPONSE TRANSCRIPTION FACTORs* (*ARFs*), were reduced upon heat stress (Kruszka et al., 2014). In *Arabidopsis*, the miR398 expression was rapidly induced by heat stress, while its target genes like *CSD* (encoding the copper/zinc SOD) and *CCS* (encoding a chaperone for CSD) were downregulated by heat stress (Guan et al., 2013). Moreover, heat shock factors HSFA1b and HSFA7b are required for heat stress induced the miR398 expression. Furthermore, the heat stress-induced miR156 plays a crucial role in regulating heat stress memory *via* repressing expression of *SPL* (*SQUAMOSA-PROMOTER BINDING-LIKE*) genes (Stief et al., 2014). These results suggest that miRNAs are generally important for heat stress response in plants.

In both heat-tolerant and heat-sensitive cotton anthers, heat stress repressed the miR156 expression, which consequently increased expression of its target *SPL* genes (Ding et al., 2017). The miR160 expression was suppressed in heat-tolerant cotton

but increased in heat-sensitive cotton under heat stress. MiR160 target genes *ARF10* and *ARF17* showed opposite expression pattens to miR160. A recent study identified a plethora of miRNAs which respond to heat stress at a stage-specific manner during cotton anther development (Chen et al., 2020). For instance, expression of miR160, miR167, and miR2949 was elevated at the sporogenous cell proliferation stage under high temperature, while miR156 responded to heat stress at male meiosis and microspore release stages. MiRNAs are also involved in epigenetic regulation *via* controlling DNA methylation and histone modification. MiR165/166 mediates methylation of downstream coding sequences of their target genes *PHABULOSA* and *PHAVOLUTA* in *Arabidopsis* (Bao et al., 2004). MiR156 and its target genes *SPLs* control transition from juvenile to adult phase in *Arabidopsis* (Xu et al., 2018; Manuela and Xu, 2020). *MIR156A* and *MIR156C* loci are major contributors to the formation of mature miR156. The H2A histone variant H2A.Z promotes expression of *MIR156A* and *MIR156C via* increasing the H3K4me3 level in these two loci (Xu et al., 2018). Although lacking direct evidence, it is possible that miRNAs cope with heat stress *via* epigenetic regulation during male reproduction in plants.

CONCLUSION AND PERSPECTIVES

Male reproductive development is highly susceptible to episodes of heat stress. Heat stress leads to impaired tapetum, abnormal microspores, and pollen abortion, which cause male sterility in plants and adversely affect yield due to failure or reduction in fertilization. Different plants respond to heat stress differently, which makes it important to identify key stages susceptible to heat stress during male reproduction. This can help take correct measures to protect plants against heat stress at specific stages during plant male reproduction.

At the molecular level, plants respond to heat stress in multiple ways. Molecular genetics, transcriptomic, and proteomic studies identified a wide array of genes and gene networks associated with heat stress during male reproduction in various crops (Giorno et al., 2013; Zhang et al., 2017; Keller and Simm, 2018; Begcy et al., 2019; Liu et al., 2020; Lohani et al., 2020; Chaturvedi et al., 2021). During male reproduction, heat stress not only affects expression of genes controlling epigenetic modifications, but also ultimately alters DNA methylation status. LncRNA and miRNA also appear important for heat stress response during plant male reproductive development, further suggesting that epigenetic control is a critical means for plants to cope with heat stress.

It is imperative to elucidate functional significance of epigenetic modifications and associated genes in heat stress response during male reproduction in economic plants. Tapetal cells, male meiocytes (microsporocytes), microspores, and pollen are sensitive to high temperature (**Figure 1**). Tapetal cells are special in terms of their endopolyploidy, formation of unique organelles (i.e., elaioplast, tapetosome, and ubisch body), highly active carbohydrate and lipid metabolism, and PCD. Tapetal cells are required for releasing haploid microspores from tetrads and for supplying energy and materials for pollen development and pollen coat formation. Numerous studies using various plants have found that heat stress affects tapetal cell differentiation and degeneration, which consequently leads to abnormal microspores and pollen abortion. Thus, it is necessary to preform single-cell transcriptomic and proteomic analyses to identify genes, gene networks, as well as particularly DNA methylation and histone modification marks that are responsible for heat stress in tapetal cells, male meiocytes, microspores, and pollen. In addition, it would be worthwhile to investigate transgenerational epigenetic effects (epigenetic memory) on heat tolerance during male reproduction in plants. CRISPR-based targeted modification of epigenetic marks has emerged as a powerful tool for improving plant traits, such as heat tolerance (Ghoshal et al., 2021). Although emerging evidence suggests the importance of epigenetic regulation for heat stress response especially during male reproduction, applying the related findings to generating thermotolerant crops *via* genetic engineering and molecular breeding is still a challenge.

AUTHOR CONTRIBUTIONS

SM and DZ conceived the idea and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The research in the DZ lab is supported by the National Science Foundation (NSF IOS-1322796) and USDA National Institute of Food and Agriculture (NIFA, 2022-67013-36294).

ACKNOWLEDGMENTS

DZ also gratefully acknowledges supports by the Shaw Scientist Award from the Greater Milwaukee Foundation, the Research and Creative Activities Support (RACAS) program at the University of Wisconsin-Milwaukee, the Bradley Catalyst Award from the UWM Research Foundation, and WiSys and UW System applied research funding programs.

REFERENCES

Abiko, M., Akibayashi, K., Sakata, T., Kimura, M., Kihara, M., Itoh, K., et al. (2005). High-temperature induction of male sterility during barley (*Hordeum vulgare* L.) anther development is mediated by transcriptional inhibition. *Sex. Plant Reprod.* 18, 91–100. doi: 10.1007/s00497-005-0004-2

Ahmed, F. E., Hall, A. E., and DeMason, D. A. (1992). Heat injury during floral development in cowpea (*Vigna unguiculata*, Fabaceae). *Am. J. Bot.* 79, 784–791. doi: 10.1002/j.1537-2197.1992.tb13655.x

Bao, N., Lye, K. W., and Barton, M. K. (2004). MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev. Cell* 7, 653–662. doi: 10.1016/j.devcel.2004.10.003

Bartee, L., Malagnac, F., and Bender, J. (2001). *Arabidopsis cmt3* chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev.* 15, 1753–1758. doi: 10.1101/gad.905701

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi: 10.1016/S0092-8674(04)00045-5

Begcy, K., Nosenko, T., Zhou, L.-Z., Fragner, L., Weckwerth, W., and Dresselhaus, T. (2019). Male sterility in maize after transient heat stress during the tetrad stage of pollen development. *Plant Physiol.* 181, 683–700. doi: 10.1104/pp.19.00707

Bhatia, G., Singh, A., Verma, D., Sharma, S., and Singh, K. (2020). Genome-wide investigation of regulatory roles of lncRNAs in response to heat and drought stress in *Brassica juncea* (Indian mustard). *Environ. Exp. Bot.* 171:103922. doi: 10.1016/j.envexpbot.2019.103922

Bita, C. E., Zenoni, S., Vriezen, W. H., Mariani, C., Pezzotti, M., and Gerats, T. (2011). Temperature stress differentially modulates transcription in meiotic anthers of heat-tolerant and heat-sensitive tomato plants. *BMC Genomics* 12:384. doi: 10.1186/1471-2164-12-384

Borges, F., Donoghue, M. T. A., LeBlanc, C., Wear, E. E., Tanurdžić, M., Berube, B., et al. (2021). Loss of small-RNA directed DNA methylation in the plant cell cycle promotes germline reprogramming and somaclonal variation. *Curr. Biol.* 31, 591–600.e594. doi: 10.1016/j.cub.2020.10.098

Browne, R. G., Li, S. F., Iacuone, S., Dolferus, R., and Parish, R. W. (2021). Differential responses of anthers of stress tolerant and sensitive wheat cultivars to high temperature stress. *Planta* 254:4. doi: 10.1007/s00425-021-03656-7

Calarco, J. P., Borges, F., Donoghue, M. T. A., Van Ex, F., Jullien, P. E., Lopes, T., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance *via* small RNA. *Cell* 151, 194–205. doi: 10.1016/j.cell.2012.09.001

Cecchetti, V., Altamura, M. M., Brunetti, P., Petrocelli, V., Falasca, G., Ljung, K., et al. (2013). Auxin controls *Arabidopsis* anther dehiscence by regulating endothecium lignification and jasmonic acid biosynthesis. *Plant J.* 74, 411–422. doi: 10.1111/tpj.12130

Chaturvedi, P., Wiese, A. J., Ghatak, A., Záveská Drábková, L., Weckwerth, W., and Honys, D. (2021). Heat stress response mechanisms in pollen development. *New Phytol.* 231, 571–585. doi: 10.1111/nph.17380

Chen, Y., Müller, F., Rieu, I., and Winter, P. (2016). Epigenetic events in plant male germ cell heat stress responses. *Plant Reprod.* 29, 21–29. doi: 10.1007/s00497-015-0271-5

Chen, J., Pan, A., He, S., Su, P., Yuan, X., Zhu, S., et al. (2020). Different MicroRNA families involved in regulating high temperature stress response during cotton (*Gossypium hirsutum* L.) anther development. *Int. J. Mol. Sci.* 21:1280. doi: 10.3390/ijms21041280

Clément, C., and Pacini, E. (2001). Anther plastids in angiosperms. *Bot. Rev.* 67:54. doi: 10.1007/BF02857849

Cross, R., McKay, S. G., McHughen, A., and Bonham-Smith, P. (2003). Heat-stress effects on reproduction and seed set in *Linum usitatissimum* L.(flax). *Plant Cell Environ.* 26, 1013–1020. doi: 10.1046/j.1365-3040.2003.01006.x

De Storme, N., and Geelen, D. (2014). The impact of environmental stress on male reproductive development in plants: biological processes and molecular mechanisms. *Plant Cell Environ.* 37, 1–18. doi: 10.1111/pce.12142

De Storme, N., and Geelen, D. (2020). High temperatures alter cross-over distribution and induce male meiotic restitution in *Arabidopsis thaliana*. *Commun. Biol.* 3:187. doi: 10.1038/s42003-020-0897-1

Ding, J., Lu, Q., Ouyang, Y., Mao, H., Zhang, P., Yao, J., et al. (2012). A long noncoding RNA regulates photoperiod-sensitive male sterility, an essential component of hybrid rice. *Proc. Natl. Acad. Sci.* 109, 2654–2659. doi: 10.1073/pnas.1121374109

Ding, Y., Ma, Y., Liu, N., Xu, J., Hu, Q., Li, Y., et al. (2017). microRNAs involved in auxin signalling modulate male sterility under high-temperature stress in cotton (*Gossypium hirsutum*). *Plant J.* 91, 977–994. doi: 10.1111/tpj.13620

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

Endo, M., Tsuchiya, T., Hamada, K., Kawamura, S., Yano, K., Ohshima, M., et al. (2009). High temperatures cause male sterility in rice plants with transcriptional alterations during pollen development. *Plant Cell Physiol.* 50, 1911–1922. doi: 10.1093/pcp/pcp135

Erickson, A., and Markhart, A. (2002). Flower developmental stage and organ sensitivity of bell pepper (*Capsicum annuum* L.) to elevated temperature. *Plant Cell Environ.* 25, 123–130. doi: 10.1046/j.0016-8025.2001.00807.x

Feng, X., Zilberman, D., and Dickinson, H. (2013). A conversation across generations: Soma-germ cell crosstalk in plants. *Dev. Cell* 24, 215–225. doi: 10.1016/j.devcel.2013.01.014

Fragkostefanakis, S., Mesihovic, A., Simm, S., Paupière, M. J., Hu, Y., Paul, P., et al. (2016). HsfA2 controls the activity of developmentally and stress-regulated heat stress protection mechanisms in tomato male reproductive tissues. *Plant Physiol.* 170, 2461–2477. doi: 10.1104/pp.15.01913

Fragkostefanakis, S., Simm, S., Paul, P., Bublak, D., Scharf, K. D., and Schleiff, E. (2015). Chaperone network composition in *Solanum lycopersicum* explored by transcriptome profiling and microarray meta-analysis. *Plant Cell Environ.* 38, 693–709. doi: 10.1111/pce.12426

Frank, G., Pressman, E., Ophir, R., Althan, L., Shaked, R., Freedman, M., et al. (2009). Transcriptional profiling of maturing tomato (*Solanum lycopersicum* L.) microspores reveals the involvement of heat shock proteins, ROS scavengers, hormones, and sugars in the heat stress response. *J. Exp. Bot.* 60, 3891–3908. doi: 10.1093/jxb/erp234

Gao, G., Li, J., Li, H., Li, F., Xu, K., Yan, G., et al. (2014). Comparison of the heat stress induced variations in DNA methylation between heat-tolerant and heat-sensitive rapeseed seedlings. *Breed. Sci.* 64, 125–133. doi: 10.1270/jsbbs.64.125

Ghoshal, B., Picard, C. L., Vong, B., Feng, S., and Jacobsen, S. E. (2021). CRISPR-based targeting of DNA methylation in *Arabidopsis thaliana* by a bacterial CG-specific DNA methyltransferase. *Proc. Natl. Acad. Sci.* 118:e2125016118. doi: 10.1073/pnas.2125016118

Giorno, F., Wolters-Arts, M., Mariani, C., and Rieu, I. (2013). Ensuring reproduction at high temperatures: the heat stress response during anther and pollen development. *Plan. Theory* 2, 489–506. doi: 10.3390/plants2030489

Goldberg, R. B., Beals, T. P., and Sanders, P. M. (1993). Anther development: basic principles and practical applications. *Plant Cell* 5, 1217–1229. doi: 10.1105/tpc.5.10.1217

González-Schain, N., Dreni, L., Lawas, L. M., Galbiati, M., Colombo, L., Heuer, S., et al. (2016). Genome-wide transcriptome analysis during anthesis reveals new insights into the molecular basis of heat stress responses in tolerant and sensitive rice varieties. *Plant Cell Physiol.* 57, 57–68. doi: 10.1093/pcp/pcv174

Gross, Y., and Kigel, J. (1994). Differential sensitivity to high temperature of stages in the reproductive development of common bean (*Phaseolus vulgaris* L.). *Field Crop Res.* 36, 201–212. doi: 10.1016/0378-4290(94)90112-0

Guan, Q., Lu, X., Zeng, H., Zhang, Y., and Zhu, J. (2013). Heat stress induction of miR398 triggers a regulatory loop that is critical for thermotolerance in *Arabidopsis*. *Plant J.* 74, 840–851. doi: 10.1111/tpj.12169

Harsant, J., Pavlovic, L., Chiu, G., Sultmanis, S., and Sage, T. L. (2013). High temperature stress and its effect on pollen development and morphological components of harvest index in the C3 model grass *Brachypodium distachyon*. *J. Exp. Bot.* 64, 2971–2983. doi: 10.1093/jxb/ert142

He, J., Jiang, Z., Gao, L., You, C., Ma, X., Wang, X., et al. (2019). Genome-wide transcript and small RNA profiling reveals transcriptomic responses to heat stress. *Plant Physiol.* 181, 609–629. doi: 10.1104/pp.19.00403

Hedhly, A., Nestorova, A., Herrmann, A., and Grossniklaus, U. (2020). Acute heat stress during. Stamen development affects both the germline and sporophytic lineages in *Arabidopsis thaliana* (L.) Heynh. *Environ. Exp. Bot.* 173:103992. doi: 10.1016/j.envexpbot.2020.103992

Higashitani, A. (2013). High temperature injury and auxin biosynthesis in microsporogenesis. *Front. Plant Sci.* 4:47. doi: 10.3389/fpls.2013.00047

Hossain, M. S., Kawakatsu, T., Kim, K. D., Zhang, N., Nguyen, C. T., Khan, S. M., et al. (2017). Divergent cytosine DNA methylation patterns in single-cell, soybean root hairs. *New Phytol.* 214, 808–819. doi: 10.1111/nph.14421

Hsieh, P.-H., He, S., Buttress, T., Gao, H., Couchman, M., Fischer, R. L., et al. (2016). *Arabidopsis* male sexual lineage exhibits more robust maintenance of CG methylation than somatic tissues. *Proc. Natl. Acad. Sci.* 113, 15132–15137. doi: 10.1073/pnas.1619074114

Hsieh, K., and Huang, A. H. C. (2007). Tapetosomes in *brassica* Tapetum accumulate endoplasmic reticulum–derived flavonoids and alkanes for delivery to the pollen surface. *Plant Cell* 19, 582–596. doi: 10.1105/tpc.106.049049

Huang, J., Li, Z., Biener, G., Xiong, E., Malik, S., Eaton, N., et al. (2017). Carbonic anhydrases function in anther cell differentiation downstream of

the receptor-Like kinase EMS1. *Plant Cell* 29, 1335–1356. doi: 10.1105/tpc.16.00484

Huang, J., Li, Z., and Zhao, D. (2016). Deregulation of the OsmiR160 target gene *OsARF18* causes growth and developmental defects with an alteration of auxin signaling in rice. *Sci. Rep.* 6:29938. doi: 10.1038/srep29938

Ishiguro, S., Nishimori, Y., Yamada, M., Saito, H., Suzuki, T., Nakagawa, T., et al. (2010). The *Arabidopsis FLAKY POLLEN1* gene encodes a 3-hydroxy-3-methylglutaryl-coenzyme A synthase required for development of tapetum-specific organelles and fertility of pollen grains. *Plant Cell Physiol.* 51, 896–911. doi: 10.1093/pcp/pcq068

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., and Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature* 472, 115–119. doi: 10.1038/nature09861

Jain, M., Prasad, P. V. V., Boote, K. J., Hartwell, A. L., and Chourey, P. S. (2007). Effects of season-long high temperature growth conditions on sugar-to-starch metabolism in developing microspores of grain sorghum (*Sorghum bicolor* L. Moench). *Planta* 227, 67–79. doi: 10.1007/s00425-007-0595-y

Keller, M., and Simm, S. (2018). The coupling of transcriptome and proteome adaptation during development and heat stress response of tomato pollen. *BMC Genomics* 19:447. doi: 10.1186/s12864-018-4824-5

Korotko, U., Chwiałkowska, K., Sańko-Sawczenko, I., and Kwasniewski, M. (2021). DNA demethylation in response to heat stress in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* 22:1555. doi: 10.3390/ijms22041555

Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., and Scharf, K.-D. (2007). Complexity of the heat stress response in plants. *Curr. Opin. Plant Biol.* 10, 310–316. doi: 10.1016/j.pbi.2007.04.011

Kruszka, K., Pacak, A., Swida-Barteczka, A., Nuc, P., Alaba, S., Wroblewska, Z., et al. (2014). Transcriptionally and post-transcriptionally regulated microRNAs in heat stress response in barley. *J. Exp. Bot.* 65, 6123–6135. doi: 10.1093/jxb/eru353

Ku, S., Yoon, H., Suh, H. S., and Chung, Y.-Y. (2003). Male-sterility of thermosensitive genic male-sterile rice is associated with premature programmed cell death of the tapetum. *Planta* 217, 559–565. doi: 10.1007/s00425-003-1030-7

Lamke, J., Brzezinka, K., Altmann, S., and Bäurle, I. (2016). A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. *EMBO J.* 35, 162–175. doi: 10.15252/embj.201592593

Law, J. A., and Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220. doi: 10.1038/nrg2719

Lee, Y.-S., Maple, R., Dürr, J., Dawson, A., Tamim, S., del Genio, C., et al. (2021). A transposon surveillance mechanism that safeguards plant male fertility during stress. *Nat. Plants* 7, 34–41. doi: 10.1038/s41477-020-00818-5

Lei, X., Ning, Y., Eid Elesawi, I., Yang, K., Chen, C., Wang, C., et al. (2020). Heat stress. Interferes with chromosome segregation and cytokinesis during male meiosis in *Arabidopsis thaliana*. *Plant Signal. Behav.* 15:1746985. doi: 10.1080/15592324.2020.1746985

Li, J., Huang, Q., Sun, M., Zhang, T., Li, H., Chen, B., et al. (2016). Global DNA methylation variations after short-term heat shock treatment in cultured microspores of *Brassica napus* cv. Topas. *Sci. Rep.* 6:38401. doi: 10.1038/srep3840

Liu, J., Feng, L., Gu, X., Deng, X., Qiu, Q., Li, Q., et al. (2019). An H3K27me3 demethylase-HSFA2 regulatory loop orchestrates transgenerational thermomemory in *Arabidopsis*. *Cell Res.* 29, 379–390. doi: 10.1038/s41422-019-0145-8

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

Liu, G., Zha, Z., Cai, H., Qin, D., Jia, H., Liu, C., et al. (2020). Dynamic transcriptome analysis of anther response to heat stress during anthesis in thermotolerant rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* 21:1155. doi: 10.3390/ijms21031155

Lohani, N., Singh, M. B., and Bhalla, P. L. (2020). High temperature susceptibility of sexual reproduction in crop plants. *J. Exp. Bot.* 71, 555–568. doi: 10.1093/jxb/erz426

Lv, Y., Hu, F., Zhou, Y., Wu, F., and Gaut, B. S. (2019). Maize transposable elements contribute to long non-coding RNAs that are regulatory hubs for abiotic stress response. *BMC Genomics* 20:864. doi: 10.1186/s12864-019-6245-5

Ma, Y., Min, L., Wang, M., Wang, C., Zhao, Y., Li, Y., et al. (2018). Disrupted genome methylation in response to high temperature has distinct affects on microspore abortion and anther indehiscence. *Plant Cell* 30, 1387–1403. doi: 10.1105/tpc.18.00074

Mangelsen, E., Kilian, J., Harter, K., Jansson, C., Wanke, D., and Sundberg, E. (2011). Transcriptome analysis of high-temperature stress in developing barley caryopses: early stress responses and effects on storage compound biosynthesis. *Mol. Plant* 4, 97–115. doi: 10.1093/mp/ssq058

Manuela, D., and Xu, M. (2020). Juvenile leaves or adult leaves: determinants for vegetative phase change in flowering plants. *Int. J. Mol. Sci.* 21:9753. doi: 10.3390/ijms21249753

Mariani, C., Beuckeleer, M. D., Truettner, J., Leemans, J., and Goldberg, R. B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. *Nature* 347, 737–741. doi: 10.1038/347737a0

Min, L., Li, Y. Y., Hu, Q., Zhu, L. F., Gao, W. H., Wu, Y. L., et al. (2014). Sugar and auxin signaling pathways respond to high-temperature stress during anther development as revealed by transcript profiling analysis in cotton. *Plant Physiol.* 164, 1293–1308. doi: 10.1104/pp.113.232314

Murphy, K. M., Egger, R. L., and Walbot, V. (2015). Chloroplasts in anther endothecium of *Zea mays* (Poaceae). *Am. J. Bot.* 102, 1931–1937. doi: 10.3732/ajb.1500384

Naydenov, M., Baev, V., Apostolova, E., Gospodinova, N., Sablok, G., Gozmanova, M., et al. (2015). High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in *Arabidopsis*. *Plant Physiol. Biochem.* 87, 102–108. doi: 10.1016/j.plaphy.2014.12.022

Ohama, N., Sato, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2017). Transcriptional regulatory network of plant heat stress response. *Trends Plant Sci.* 22, 53–65. doi: 10.1016/j.tplants.2016.08.015

Omidi, M., Siahpoosh, M. R., Mamghani, R., and Modarresi, M. (2014). The influence of terminal heat stress on meiosis abnormalities in pollen mother cells of wheat. *Cytologia* 79, 49–58. doi: 10.1508/cytologia.79.49

Oshino, T., Abiko, M., Saito, R., Ichiishi, E., Endo, M., Kawagishi-Kobayashi, M., et al. (2007). Premature progression of anther early developmental programs accompanied by comprehensive alterations in transcription during high-temperature injury in barley plants. *Mol. Gen. Genomics.* 278, 31–42. doi: 10.1007/s00438-007-0229-x

Pacini, E., Franchi, G. G., and Hesse, M. (1985). The Tapetum: its form, function, and possible phylogeny in Embryophyta. *Plant Syst. Evol.* 149, 155–185. doi: 10.1007/BF00983304

Papareddy, R. K., and Nodine, M. D. (2021). Plant epigenetics: propelling DNA methylation variation across the cell cycle. *Curr. Biol.* 31, R129–R131. doi: 10.1016/j.cub.2020.11.049

Parish, R. W., and Li, S. F. (2010). Death of a tapetum: a programme of developmental altruism. *Plant Sci.* 178, 73–89. doi: 10.1016/j.plantsci.2009.11.001

Parish, R. W., Phan, H. A., Iacuone, S., and Li, S. F. (2012). Tapetal development and abiotic stress: a centre of vulnerability. *Funct. Plant Biol.* 39, 553–559. doi: 10.1071/FP12090

Pecinka, A., Dinh, H. Q., Baubec, T., Rosa, M., Lettner, N., and Scheid, O. M. (2010). Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. *Plant Cell* 22, 3118–3129. doi: 10.1105/tpc.110.078493

Popova, O. V., Dinh, H. Q., Aufsatz, W., and Jonak, C. (2013). The RdDM pathway is required for basal heat tolerance in *Arabidopsis*. *Mol. Plant* 6, 396–410. doi: 10.1093/mp/sst023

Prasad, P. V., Boote, K. J., Allen, L. H. Jr., and Thomas, J. M. (2002). Effects of elevated temperature and carbon dioxide on seed-set and yield of kidney bean (*Phaseolus vulgaris* L.). *Glob. Chang. Biol.* 8, 710–721. doi: 10.1046/j.1365-2486.2002.00508.x

Pressman, E., Peet, M. M., and Pharr, D. M. (2002). The effect of heat stress on tomato pollen characteristics is associated with changes in carbohydrate concentration in the developing anthers. *Ann. Bot.* 90, 631–636. doi: 10.1093/aob/mcf240

Qian, Y., Hu, W., Liao, J., Zhang, J., and Ren, Q. (2019a). The dynamics of DNA methylation in the maize (*Zea mays* L.) inbred line B73 response to heat stress at the seedling stage. *Biochem. Biophys. Res. Commun.* 512, 742–749. doi: 10.1016/j.bbrc.2019.03.150

Qian, Y., Ren, Q., Zhang, J., and Chen, L. (2019b). Transcriptomic analysis of the maize (*Zea mays* L.) inbred line B73 response to heat stress at the seedling stage. *Gene* 692, 68–78. doi: 10.1016/j.gene.2018.12.062

Rutley, N., Poidevin, L., Doniger, T., Tillett, R. L., Rath, A., Forment, J., et al. (2021). Characterization of novel pollen-expressed transcripts reveals their potential roles in pollen heat stress response in *Arabidopsis thaliana*. *Plant Reprod.* 34, 61–78. doi: 10.1007/s00497-020-00400-1

Saini, H., and Aspinall, D. (1982). Abnormal sporogenesis in wheat (*Triticum aestivum* L.) induced by short periods of high temperature. *Ann. Bot.* 49, 835–846. doi: 10.1093/oxfordjournals.aob.a086310

Saini, H., Sedgley, M., and Aspinall, D. (1984). Development anatomy in wheat of male sterility induced by heat stress, water deficit or abscisic acid. *Funct. Plant Biol.* 11, 243–253. doi: 10.1071/PP9840243

Sakata, T., Oshino, T., Miura, S., Tomabechi, M., Tsunaga, Y., Higashitani, N., et al. (2010). Auxins reverse plant male sterility caused by high temperatures. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8569–8574. doi: 10.1073/pnas.1000869107

Sanders, P. M., Bui, A. Q., Weterings, K., McIntire, K. N., Hsu, Y.-C., Lee, P. Y., et al. (1999). Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sex. Plant Reprod.* 11, 297–322. doi: 10.1007/s004970050158

Sato, H., Mizoi, J., Tanaka, H., Maruyama, K., Qin, F., Osakabe, Y., et al. (2014). *Arabidopsis* DPB3-1, a DREB2A interactor, specifically enhances heat stress-induced gene expression by forming a heat stress-specific transcriptional complex with NF-Y subunits. *Plant Cell* 26, 4954–4973. doi: 10.1105/tpc.114.132928

Sato, H., Suzuki, T., Takahashi, F., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2019). NF-YB2 and NF-YB3 have functionally diverged and differentially induce drought and heat stress-specific genes. *Plant Physiol.* 180, 1677–1690. doi: 10.1104/pp.19.00391

Scott, R. J., Spielman, M., and Dickinson, H. G. (2004). Stamen structure and function. *Plant Cell* 16, S46–S60. doi: 10.1105/tpc.017012

Slotkin, R. K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J. D., Feijó, J. A., et al. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* 136, 461–472. doi: 10.1016/j.cell.2008.12.038

Smith, A. R., and Zhao, D. Z. (2016). Sterility caused by floral organ degeneration and abiotic stresses in *Arabidopsis* and cereal grains. *Front. Plant Sci.* 7:1503. doi: 10.3389/fpls.2016.01503

Song, Y., Chen, Q., Ci, D., Shao, X., and Zhang, D. (2014). Effects of high temperature on photosynthesis and related gene expression in poplar. *BMC Plant Biol.* 14:111. doi: 10.1186/1471-2229-14-111

Song, G., Wang, M., Zeng, B., Zhang, J., Jiang, C., Hu, Q., et al. (2015). Anther response to high-temperature stress during development and pollen thermotolerance heterosis as revealed by pollen tube growth and in vitro pollen vigor analysis in upland cotton. *Planta* 241, 1271–1285. doi: 10.1007/s00425-015-2259-7

Stief, A., Altmann, S., Hoffmann, K., Pant, B. D., Scheible, W.-R., and Bäurle, I. (2014). *Arabidopsis miR156* regulates tolerance to recurring environmental stress through SPL transcription factors. *Plant Cell* 26, 1792–1807. doi: 10.1105/tpc.114.123851

Suzuki, K., Takeda, H., Tsukaguchi, T., and Egawa, Y. (2001). Ultrastructural study on degeneration of tapetum in anther of snap bean (*Phaseolus vulgaris* L.) under heat stress. *Sex. Plant Reprod.* 13, 293–299. doi: 10.1007/s004970100071

Vara Prasad, P. V., Craufurd, P. Q., and Summerfield, R. J. (1999). Fruit number in relation to pollen production and viability in groundnut exposed to short episodes of heat stress. *Ann. Bot.* 84, 381–386. doi: 10.1006/anbo.1999.0926

Walbot, V., and Egger, R. L. (2016). Pre-meiotic anther development: cell fate specification and differentiation. *Annu. Rev. Plant Biol.* 67, 365–395. doi: 10.1146/annurev-arplant-043015-111804

Wang, A., Hu, J., Gao, C., Chen, G., Wang, B., Lin, C., et al. (2019). Genome-wide analysis of long non-coding RNAs unveils the regulatory roles in the heat tolerance of Chinese cabbage (*Brassica rapa* ssp.chinensis). *Sci. Rep.* 9:5002. doi: 10.1038/s41598-019-41428-2

Wang, A., Xia, Q., Xie, W., Datla, R., and Selvaraj, G. (2003). The classical Ubisch bodies carry a sporophytically produced structural protein (RAFTIN) that is essential for pollen development. *Proc. Natl. Acad. Sci.* 100, 14487–14492. doi: 10.1073/pnas.2231254100

Wu, S. S., Platt, K. A., Ratnayake, C., Wang, T. W., Ting, J. T., and Huang, A. H. (1997). Isolation and characterization of neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12711–12716. doi: 10.1073/pnas.94.23.12711

Xin, M., Wang, Y., Yao, Y., Song, N., Hu, Z., Qin, D., et al. (2011). Identification and characterization of wheat long non-protein coding RNAs responsive to powdery mildew infection and heat stress by using microarray analysis and SBS sequencing. *BMC Plant Biol.* 11:61. doi: 10.11861471-2229-11-61

Xu, M., Leichty, A. R., Hu, T., and Poethig, R. S. (2018). H2A.Z promotes the transcription of *MIR156A* and *MIR156C* in *Arabidopsis* by facilitating the deposition of H3K4me3. *Development* 145:dev152868. doi: 10.1242/dev.152868

Yamaguchi, N., Matsubara, S., Yoshimizu, K., Seki, M., Hamada, K., Kamitani, M., et al. (2021). H3K27me3 demethylases alter *HSP22* and *HSP17.6C* expression in response to recurring heat in *Arabidopsis*. *Nat. Commun.* 12:3480. doi: 10.1038/s41467-021-23766-w

Yamakawa, H., Hirose, T., Kuroda, M., and Yamaguchi, T. (2007). Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiol.* 144, 258–277. doi: 10.1104/pp.107.098665

Yang, K. A., Lim, C. J., Hong, J. K., Park, C. Y., Cheong, Y. H., Chung, W. S., et al. (2006). Identification of cell wall genes modified by a permissive high temperature in Chinese cabbage. *Plant Sci.* 171, 175–182. doi: 10.1016/j.plantsci.2006.03.013

Yu, X., Wang, H., Lu, Y., de Ruiter, M., Cariaso, M., Prins, M., et al. (2011). Identification of conserved and novel microRNAs that are responsive to heat stress in *Brassica rapa*. *J. Exp. Bot.* 63, 1025–1038. doi: 10.1093/jxb/err337

Yu, Y., Zhang, Y., Chen, X., and Chen, Y. (2019). Plant noncoding RNAs: hidden players in development and stress responses. *Annu. Rev. Cell Dev. Biol.* 35, 407–431. doi: 10.1146/annurev-cellbio-100818-125218

Zhang, D., Liu, D., Lv, X., Wang, Y., Xun, Z., Liu, Z., et al. (2014). The cysteine protease CEP1, a key executor involved in tapetal programmed cell death, regulates pollen development in Arabidopsis. *Plant Cell* 26, 2939–2961. doi: 10.1105/tpc.114.127282

Zhang, S.-S., Yang, H., Ding, L., Song, Z.-T., Ma, H., Chang, F., et al. (2017). Tissue-specific transcriptomics reveals an important role of the unfolded protein response in maintaining fertility upon heat stress in Arabidopsis. *Plant Cell* 29, 1007–1023. doi: 10.1105/tpc.16.00916

Zhao, D. (2009). Control of anther cell differentiation: a teamwork of receptor-like kinases. *Sex. Plant Reprod.* 22, 221–228. doi: 10.1007/s00497-009-0106-3

Zhao, D. Z., Wang, G. F., Speal, B., and Ma, H. (2002). The *EXCESS MICROSPOROCYTES1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes Dev.* 16, 2021–2031. doi: 10.1101/gad.997902

Zhao, Q., Zhou, L., Liu, J., Cao, Z., Du, X., Huang, F., et al. (2018). Involvement of CAT in the detoxification of HT-induced ROS burst in rice anther and its relation to pollen fertility. *Plant Cell Rep.* 37, 741–757. doi: 10.1007/s00299-018-2264-y

Zoong Lwe, Z. S., Welti, R., Anco, D., Naveed, S., Rustgi, S., and Narayanan, S. (2020). Heat stress elicits remodeling in the anther lipidome of peanut. *Sci. Rep.* 10:22163. doi: 10.1038/s41598-020-78695-3

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership